THE UNIVERSITY OF CHICAGO

UNDERSTANDING CONVECTIVE AVAILABLE POTENTIAL ENERGY (CAPE) IN CURRENT AND FUTURE CLIMATES WITH SIMPLE FRAMEWORKS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF THE GEOPHYSICAL SCIENCES

BY
ZIWEI WANG

CHICAGO, ILLINOIS
JUNE 2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

As I step closer to completing my PhD degree, the more grateful I am for the support I have got, without which none of this is possible. I am most grateful to my advisor, Liz Moyer. I view her as not only a mentor but also a friend. I have learned so much from Liz, starting with building physical intuitions, identifying fascinating scientific questions, creating a pipeline for research, testing a hypothesis, to finally presenting my ideas. These skills have become part of me, and they will be an invaluable life-long asset. I sincerely thank Dan Chavas, Tiffany Shaw, Amir Jina, and Noboru Nakamura for their scientific insight and valuable guidance, which has helped me grow in multiple directions.

I thank my parents, Baomin Wang, and Yan Guo, for their support and sharing their wisdom about life.

I am lucky to have met amazing mentors and collaborators in the Data Science for Energy and Environmental Research program, where I could explore my broad interests at the intersection between climate science and public policy, statistics, and computer science. I thank faculty members from Peking Unversity and UCLA, who helped me build a fundamental basis and develop a deep interest in Atmospheric Science; I feel especially honored to have the chance to pursue my Doctoral degree at the University of Chicago, where I learned to understand the nature and world with simple models.

It's been an extremely joyful and rewarding journey towards completing this thesis. I want to thank all my friends for making my life here even more unforgettable. The past five and a half years have been transformative for my personal growth, and I feel blessed with the extraordinary skills that I have obtained along the way. Now I can see the world differently.

Lastly, I acknowledge the support from the National Science Foundation (SES-1463644) and the University of Chicago Research Computing Center for providing the computational resources that made this work possible.

STATEMENT OF ORIGINALITY AND COPYRIGHT ATTRIBUTION

All text and analysis in Chapters 2 and 3 was produced in collaboration with the authors listed at the beginning of each chapter, with Ziwei Wang serving as the lead author in each case. All figures contained within were produced by Ziwei Wang using data from the publicly available sources listed in each section. All figures and text in Chapters 2 and 3 have been published at the following locations:

- **Wang, Z.,** J. A. Franke, Z. Luo, and E. J. Moyer. Reanalyses and a High-Resolution Model Fail to Capture the "High Tail" of CAPE Distributions. *Journal of Climate* (2021). https://doi.org/10.1175/JCLI-D-20-0278.1. **© American Meteorological Society. Used with permission.** American Meteorological Society holds the rights but grants permission to authors for full reproduction for their dissertation.

- **Wang, Z.,** and E. J. Moyer. Robust Relationship between Midlatitudes CAPE and Moist Static Energy Surplus in Present and Future Simulations. *Geophysical Research Letters* (2023). https://doi.org/10.1029/2023GL104163. Published under Creative Commons Attribution-NonCommercial 4.0 International copyright: `https://creativecommons.org/licenses/by-nc/4.0/`.

All text and analysis in Chapter 4 was produced by Ziwei Wang.

ABSTRACT

Importance: Convective storms are one of the most costly natural hazards in the current climate. These storms are expected to change with future climate warming and may bring more devastating socioeconomic costs. Our understanding of these storms has substantially improved with advancements in theoretical understanding, model physics representation, and computational power. However, combining insights from simplified frameworks with those from realistic representations of these storms at scale remains challenging. Large-scale metrics for convective storms, like Convective Available Potential Energy (CAPE), could help bridge the gap between these paths. This thesis aims: (1) to evaluate existing model and reanalyses representation of CAPE distributions, (2) to understand the primary drivers of CAPE changes between climate states, and (3) to connect weather and climate scale variations with a full scaling of CAPE.

Approach: This work uses an array of observations and modeling datasets with varying levels of authenticity and complexity. We approach the scientific questions by combining these datasets with simple theoretical frameworks. In Chapter 2, we use a radiosonde observational dataset (IGRA), against which we evaluate reanalyses (ERA-Interim and ERA5) and a convection-permitting model (WRF) in the current climate. In Chapter 3, we explain the full distributional projection of CAPE from the convection-permitting WRF simulations with CAPE-MSE surplus dependence and synthetic profiles. In Chapter 4, we propose and evaluate the full scaling of CAPE with ERA5 reanalysis and 11 models from the CMIP6 inter-comparison project across space and different temporal scales.

Key Findings: We find that warming increases the occurrence of high CAPE conditions substantially in all climate models. While CAPE distributions in coarse-resolution models are not accurate, CAPE in a high-resolution convection-permitting model largely matches observations other than in extreme CAPE conditions, whose occurrence is underestimated in the current climate. The low biases arise from an underprediction of hot and humid conditions. Future projections of midlatitudes CAPE exhibit distributional shifts, so they cannot be expressed as a simple mean change; they also cannot be sufficiently predicted by changes in surface conditions alone. We find that the distributional shift can be captured with three mean changes at both the surface (T_s, RH_s) and mid-troposphere (T_m), highlighting the importance of a lapse rate adjustment in mid-latitude summertime under climate change. Furthermore, the minimal three state parameters can be reduced to a single parameter of "MSE surplus". CAPE dependence on MSE surplus remains consistent across climate states in both the nudged convection-permitting model (WRF) and in 11 free-running CMIP6 models. On shorter timescales than climatological shifts, predicting CAPE variations requires at least one additional input, the convective layer depth. We, therefore, derived a robust CAPE scaling from entropy and buoyancy forms that effectively captures CAPE variations across spatial and temporal scales, including diurnal, seasonal, and climatological variations. This scaling provides key physical insights into how much and why CAPE changes, with strong implications for societal impacts. It allows model biases to be diagnosed and may provide a practical tool for weather analysis.

CHAPTER 1

INTRODUCTION

Convective extremes are one of the most costly natural hazards, causing substantial socio-economic damage across sectors [69, 9, 66, 39]. Among the 26 billion-dollar weather events in the Contiguous United States, 19 are related to thunderstorms; these events take on forms of heavy rainfall, lightning, hail, high wind, tornadoes, or a combination of them [120]. It is critical to understand the characteristics of the convective population, especially their intensity and occurrence, in both current and future climates [90, 19]. The most commonly used large-scale metric for convective extreme is Convective Available Potential Energy (CAPE) [86]. This thesis aims to improve our understanding of convective populations in current and future climates with simple frameworks developed for CAPE.

CAPE is the vertically integrated buoyancy for an adiabatically rising parcel [86]. It is directly linked to the maximum updraft velocity – strong updrafts are key to severe convective storms. On synoptic time scales, CAPE is used to forecast convective extreme events. CAPE is routinely reported in all major numerical weather prediction models (e.g. GFS, NAM, HRRR) at a sub-daily time scale. The covariate of CAPE and vertical wind shear is a key proxy for the occurrence of severe thunderstorms [14, 11]; the covariate of CAPE and precipitation is used to predict the occurrence of lightning [111]. The predictive power of these proxies has been consolidated in multiple papers for types of events including hail, tornadoes, and waterspouts over Europe [e.g. 58, 102, 127, 126], North America [e.g. 13, 74, 110, 45], Australia [e.g. 3, 15] and other regions [e.g. 153, 113].

An important approach to understanding convective populations is with models. In low-resolution models like global general circulation models, convection is parametrized due to the insufficient resolution to resolve fine-scale dynamics. In these schemes, the intensity and/or triggering of convection (especially deep convection) is usually parametrized based on CAPE [94, 47]. Some of the most commonly used schemes determine convective mass-flux based on CAPE-relaxing closure (CR closure), for example, Z–M scheme [149, 89], Tiedtke scheme [131], and Bechtold scheme [6]. These schemes also adopt CAPE in their triggering functions [145], and multiple works have shown that incorporating modified versions of CAPE generation rate (dCAPE) as part of the triggering function improves the representation precipitation in models [e.g. 124, 121, 122]. In high-resolution (e.g., convection-permitting) models, convection is more explicitly represented. CAPE is a key diagnostic of these models: how CAPE is accumulated and consumed has important implications for the model representation of the hydrological cycles [e.g. 104, 112, 51].

CAPE is expected to change with warming, providing important socio-economical implications. Modeling work has used the same covariate between CAPE and vertical wind shear as a proxy of severe thunderstorm conditions for future projections and found a robust increase in the occurrence of these conditions [133, 21, 10]. The model uncertainties, especially in how convective parametrization is treated in these low-resolution models, contribute substantially to the projection uncertainties for the future climate [34, 33]. More recent work leverages model ensembles (e.g. inter-comparison projects CMIP5 and CMIP6) to provide better constraints for the projected future convective extremes [134, 35, 80]. Regional convection-permitting model runs suggest that CAPE changes 6–12%/K over the tropical ocean or under RCE [108, 88, 117, 115], and 11%/K for three stations in the Eastern U.S. [104].

Previous work has proposed simple models and theories for CAPE, which usually provides a fundamental understanding of the physical processes and adds value to the existing model studies. These theoretical works fall under two major categories. One approach assumes that the strong entrainment sets the environmental profile almost completely and thus predicts CAPE based on entrainment and surface conditions alone using a “zero-buoyancy model” [117] (later expanded by [154]). The model explains how CAPE responds to surface temperature and RH in climatological mean profiles [115], and explains high CAPE climatology beyond tropical ocean [116]. [109] proposes an analytical expression for CAPE and finds that its increase follows C–C scaling under Radiative-Convective Equilibrium (RCE). However, these theories have limitations when variations on smaller spatial and shorter time scales are of interest. The other approach treats surface and mid-tropospheric conditions as independent parameters. Early attempts model the convective atmosphere as a heat engine, with entropy added at (or near) the surface through latent heat and removed aloft through radiative cooling [46, 106]. Though concerns have been raised about the efficiency term underestimating the role of dissipation as an entropy sink [109, 119], we believe treating mid-tropospheric conditions (or lapse rate) separately is necessary to capture CAPE variations at full spatial and temporal scales.

There remains a gap between theories developed under climatological mean conditions (e.g. RCE) and theories for CAPE variations on finer spatial and temporal scales. [2] first proposed a simple model for the midlatitudes transient peak of CAPE by appending a cold and dry air column onto a moist boundary layer. The simple model is further expanded to a full scaling of CAPE and validates its ability to capture the diurnal and spatial variation of CAPE over CONUS [82, LC21 hereafter]. However, the scaling in LC21 is framed around dry static energy. It thus is a less direct path to approximating CAPE than one with moist static energy (which is conserved for moist adiabatic ascent). The LC21 scaling requires deriving average properties with the full tropospheric static energy profile. Therefore, its time complexity is closer to raw CAPE calculation than a simplified scaling with only scalars. A simple scaling for CAPE variations on all spatial and temporal scales is greatly needed.

CHAPTER 2

REANALYSES AND A HIGH-RESOLUTION MODEL FAIL TO CAPTURE THE "HIGH TAIL" OF CAPE DISTRIBUTIONS

Published in *Journal of Climate* in collaboration with:

James Franke[1,2], Zhenqi Luo[3] and Elisabeth Moyer[1,2]

[1]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA [3]College of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing, China

Convective available potential energy (CAPE) is of strong interest in climate modeling because of its role in both severe weather and in model construction. Extreme levels of CAPE (> 2000 J/kg) are associated with high-impact weather events, and CAPE is widely used in convective parametrizations to help determine the strength and timing of convection. However, to date, no study has systematically evaluated CAPE biases in models in a climatological context in an assessment large enough to characterize the "high tail" of the CAPE distribution. This work compares CAPE distributions in over 200,000 summertime proximity soundings from four sources: the observational radiosonde network (IGRA), 0.125-degree reanalysis (ERA-Interim and ERA5), and a 4 km convection-permitting regional WRF simulation driven by ERA-Interim. Both reanalyses and models consistently show too-narrow distributions of CAPE, with the "high tail" ($>$ 95th percentile) systematically biased low by up to 10% in surface-based CAPE and 20% at the most unstable layer. This "missing tail" corresponds to the most impact-relevant conditions. CAPE bias in all datasets is driven by bias in surface temperature and humidity: reanalyses and model undersample observed cases of extreme heat and moisture. These results suggest that reducing inaccuracies in land surface and boundary layer models is critical for accurately reproducing CAPE.

2.1 Introduction

Convective Available Potential Energy (CAPE) is an integral quantity of buoyancy in the convective layer [86], and is considered as a key parameter in convection initiation and development. Closely linked to updraft strength and storm intensity, CAPE provides a way to understand the potential threat of some high-impact weather events such as thunderstorms, hail, and tornadoes. [14] propose a combination of CAPE and bulk wind shear as a metric for severe weather in reanalyses, with a 2000 J/kg as a threshold value for extreme events, and multiple subsequent studies confirm this relationship in models and observations. Studies relating high CAPE values to extreme precipitation or intense storms in observations [e.g. 58, 81, 36], and many others. In models, [93], for example, show that the number of extreme precipitation events in general circulation models (GCMs) grows with the covariate between CAPE and wind shear.

CAPE is also used as a key parameter in convective schemes in GCMs to determine convective mass flux [149, 146, 5]. In CAPE-relaxing closure (CR closure) schemes, modelers commonly rely on CAPE to trigger convection and to determine the total vertical mass flux so that the magnitude of vertical mass flux is directly affected by an inaccurate representation of CAPE [79, 27]. In some recently developed new schemes intended to reproduce the diurnal cycle more realistically, convective triggering is directly dependent on CAPE generation rate (dCAPE) [143, 136]. These schemes have been shown to improve model performance for precipitation diurnal peak time compared to schemes using classic CR closure [121, 142]. However, it should be noted that these dynamical-based trigger functions introduce additional sensitivity to CAPE biases, and [122] find that dCAPE trigger functions are highly sensitive to model resolution.

CAPE is derived from vertical profiles of temperature, pressure, and humidity, which are measured in situ only from a sparse network of specialized weather stations. Radiosondes measure atmospheric profiles from weather balloons released twice a day from $\sim$ 1000 stations globally. Because radiosonde measurements are both spatially and temporally sparse, researchers linking measured CAPE to severe weather events have used "proximity soundings", estimating the severity of extreme weather events based on soundings taken within a range of ∼200 km [e.g. 13, 103, 12]. More recent studies of CAPE and severe weather use not soundings but reanalyses that assimilate in-situ and remote observations in global models to provide information at higher resolution [14, 81, 36]. Global gridded reanalyses also allow ready construction of climatologies: for example, [107] use the ERA40 reanalysis to construct a 40-year climatology of CAPE, showing that the largest values and variability are found over tropical land (mean ∼2000 J/kg), with a stronger dependence on specific humidity than temperature.

To diagnose potential changes in CAPE under future higher CO_2 conditions, studies must rely on numerical simulations. With the growth of computational resources, the horizontal resolution of models used for this purpose has increased. For example, [134] and [35] examine changes in CAPE and wind shear in GCM projections (∼100 km) and infer a likely future increase in the number of days with severe weather events. [116] use both GCMs and super-parametrized GCMs (20 km) to study changes in the 95th percentile of CAPE in the tropics and subtropics during heavy precipitation and find a 6–14% increase per K regional temperature increase. (Note that CAPE values during heavy precipitation are low, [e.g. 1]; the 95th percentile in observations in [116] is under 2000 J/kg.) [104] examine changes in CAPE and convective inhibition (CIN) in a 4 km dynamically downscaled simulation of North America in a pseudo global warming scenario (driven by reanalysis or by reanalysis with an applied offset in climate variables). They find that both CAPE and CIN generally increase under warmer conditions and infer a future intensification of convective strength. Such convection-permitting models, with their improvement in convective dynamics, have been assumed to help improve the representation of CAPE.

Given the extent of the scientific use of reanalyses and model simulations, it is valuable to ask how well these products reproduce realistic CAPE values. Coarse-resolution general circulation models reproduce large-scale spatial patterns in CAPE but can produce large biases in individual locations [23]. While reanalyses and high-resolution forecast models are generally assumed to be more accurate, assessments of their biases versus radiosonde observations have been limited. Studies using restricted samples of soundings near severe weather events have produced inconsistent results. For example, [130] evaluate surface-based CAPE (SBCAPE) from the Rapid Update Cycle (RUC-2) weather prediction system 0-hour analysis against radiosondes sampled near supercells (149 soundings from 1999–2001, in the U.S. Central and Southern Plains) and find a low bias of $\sim$16% (mean bias of about -400 J/kg in mean conditions of $\sim$2500 J/kg). [25] compare SBCAPE in the RUC 0-hour analysis with a different sample of soundings near supercell thunderstorms (582 soundings during the VORTEX2 campaign in 2009–2010, also in the Central and Southern Plains) and find a small high bias ($\sim$150 J/kg) with a large spread. [4] compare mixed-layer CAPE (MLCAPE) in the reanalysis product ERA-Interim (ERAI) and in the Australian MesoLAPS (Mesoscale Limited Area Prediction System) weather model with radiosonde soundings near thunderstorm events (3697 and 4988 soundings, respectively, from 2003–2010, from 16 stations in Australia) and find slight high biases of 6 and 74 J/kg in conditions of 234 and 255 J/kg mean non-zero MLCAPE.

To date, very few validation studies have systematically evaluated CAPE bias and errors in a climatological context with a large enough scale to allow evaluation of the "high tail" of the CAPE distribution. For convection-permitting models, it is widely assumed that improved resolution also improves the representation of CAPE, but this assumption has not been explicitly tested. For reanalyses, only a few studies have compared output to large collections of soundings, and none assess distributional changes. [55] compare NARR (the North American Regional Reanalysis) to all radiosondes over 11 years from 21 stations in the Eastern U.S. (>100,000 soundings with nonzero SBCAPE from 2000–2011), but do not assess either mean bias or distributional differences. (They do find considerable spread in SBCAPE errors, with RMSE ∼1400 J/kg.) [127] and [128] use even larger sample sizes (>1 million profiles from 1979–2016, and >5 million profiles from 1980–2018, respectively, predominantly over Europe and spanning all seasons) and compare CAPE under various definitions between soundings and ERAI and ERA5 reanalyses. They examine mean biases (in soundings with non-zero CAPE) and find them large relative to median CAPE, but median values in both studies are small (∼68 and 100 J/kg for MLCAPE).

Even fewer studies have attempted to attribute bias in CAPE to specific model issues. While one possible cause is error in free tropospheric profiles, multiple authors have noted the potential role of incorrect temperature and humidity at the surface or boundary layer. Several studies have explicitly tested this attribution by replacing surface values in models and data products with observed ones and noting the improved match to radiosonde SBCAPE. [25] replaces surface values in RUC with those from the operational surface objective analysis system (SFCOA) and finds a reduction in bias in 1-hour forecasts. [54] compare 10 years of SBCAPE from a single station, the Southern Great Plains Atmospheric Radiation Measurement (ARM) site, and show that replacing surface values largely corrects CAPE values in ERAI reanalysis and values derived from the AIRS satellite. Similarly, in a very small sample (2 individual case studies), [7] find that replacing surface values of humidity and temperature corrects a low bias in SBCAPE in a satellite-derived product.

This work seeks to address both needs for large-scale systematic assessment of CAPE distributions in reanalyses and high-resolution simulations against those in radiosondes and for attribution of the source of any bias. To focus on conditions that promote strong convection, we examine summertime data over the contiguous United States, using a total of nearly 200,000 soundings over 12 years of observations.

2.2 Data Description

This study compares four datasets that allow calculation of CAPE over the contiguous United States from January 2001 to December 2012: radiosonde observations from the Integrated Global Radiosonde Archive (IGRA) version 2 [41, 42]; the reanalysis products ERA-Interim (ERAI) and ERA5 [32, 65]; and simulation output from the Weather Research and Forecasting model (WRF) at the convection-permitting resolution, forced by ERAI [105]. Because our interest is in the "high tail" of the CAPE distribution, we focus on the summer months when convection is most active and CAPE is largest. We define summer as May to August (MJJA), following the convention of many studies [e.g. 125, 104], though some work on extreme weather uses an earlier definition of April to July to include the late spring peak of convection [e.g. 134]. With this definition, IGRA provides a total of 199,787 summertime radiosonde profiles from U.S. stations with continuous records during 2001–2012. For consistency, analyses shown here involve data matched to radiosonde stations and synchronized in time, though when evaluating diurnal cycles we also show reanalysis and model output at additional times of the day.

2.2.1 Radiosonde observations

IGRA is an archive of quality-controlled atmospheric sounding profiles from weather balloons around the world collected by a standard protocol. The archive is operated by the U.S. National Oceanic and Atmospheric Administration (NOAA) and profiles in the U.S. are collected by NOAA's National Weather Service. In this work we use profiles from all stations in the contiguous United States that report continuous operation through the years 2001 to 2012, a total of 80 out of the 248 stations historically used. All stations have routine balloon launches at 00 and 12 UTC each day, though some soundings are missing (17.4% of all routine launches during this period). Many stations also include sporadic launches at 06 and 18 UTC; we include these profiles in the dataset considered here, though we generally disaggregate analyses by the time of day. Of the complete dataset of 199,787 soundings, 83,668 are from 00 UTC, 106,455 from 12 UTC, and 9,664 from additional times. All four hours are used in our analysis unless otherwise stated. Of these profiles, 1496 (0.75%) are excluded by our quality control criteria. (See Methods below.)

Variables acquired from IGRA include pressure, temperature, altitude, and vapor pressure, all of which are standard reported values. We convert vapor pressure to specific humidity and dew point temperature for consistency across all datasets. Vertical resolution varies by station, but most stations report around 80 levels from the surface to 10 hPa pressure. The data are available from `https://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive`.

2.2.2 Reanalysis products

ERAI and ERA5 are both reanalysis products maintained by the European Centre for Medium-Range Weather Forecasts (ECMWF). Both products assimilate observations into global models and are available from 1979 to the present. ERAI has a native horizontal resolution of T255 ($\approx$ 80km); it has been superseded by ERA5, which has significant improvements in spatial and temporal resolution with a native horizontal resolution of TL639 (0.28125°, $\approx$ 31km) [17]. Because our analysis involves matching individual radiosonde stations, we acquire both reanalyses at a finer spatial resolution (0.125°) produced by ECMWF with bilinear interpolation for continuous fields. We use output at native model vertical levels, preserving the highest possible vertical resolution for our CAPE calculation: 60 levels for ERAI (L60), and 137 for ERA5 (L137). We download profiles of temperature and specific humidity, and surface pressure; the pressure profile is then derived using surface pressure and coefficients a and b that define the hybrid-sigma coordinates of L60 and L137. 2m temperature and dew point temperature along with surface pressure are appended to the bottom level of profiles. Although ERA5 provides hourly output, we use data at 00, 06, 12, and 18

UTC to match with ERAI. Both products are available at `https://www.ecmwf.int/en/`.

Data assimilation is a key component of reanalysis products. Both ERAI and ERA5 assimilate a homogenized version of IGRA radiosonde observations, the Radiosonde Observation Correction using Reanalyses (RAOBCORE) [61, 62]. Reanalyses and IGRA observations are therefore not fully independent. ERAI uses a bias correction for radiosonde temperature based on RAOBCORE_T_1.3, which is further adjusted and implemented to the Continuous Observation Processing Environment (COPE) framework in ERA5 [43]. The assimilation process of ERAI uses the following exclusion criteria for radiosonde data: 1) any radiosonde observation below the model surface, and radiosonde-observed specific humidity in either 2) extreme cold conditions ($T < 193$ K for RS–90 sondes, $T < 213$ K for RS–80 sondes, $T < 233$ K otherwise), or 3) high altitude ($p < 100$ hPa for RS–80 and RS–90 sondes, $p < 300$ hPa for all other sonde types) [32].

2.2.3 High-resolution model simulation

The high-resolution model output we use is a 4-km resolution dynamically downscaled “retrospective” simulation over North America first described by [83]. The simulation is created as the control run of a pseudo-global-warming experiment and involves forcing the WRF (Weather Research and Forecasting) 3.4.1 model with ERAI reanalysis. The WRF simulation is run with 4 km grid spacing and 50 vertical levels up to 50 hPa, with parametrization schemes including: Thompson aerosol-aware microphysics [129], the Yonsei University (YSU) planetary boundary layer [68], the rapid radiative transfer model (RRTMG) [70], and the improved Noah-MP land-surface model [91].

The model uses ERAI as initial and boundary conditions, with large-scale spectral nudging applied to geopotential, temperature, and horizontal wind. Nudging is applied throughout the model domain, at all altitudes above the planetary boundary layer, and is intended to remove known large-scale issues (such as summertime high-temperature bias over the central U.S. [87]) while still allowing smaller-scale processes to modify local profiles. Values are nudged at a strength corresponding to an 'e-folding' time of 6 hours, using a wavenumber truncation of 3 and 2 in the zonal and meridional directions, respectively. Because the experiment is intended to reproduce observed snow cover over North America, some modifications are made to the land surface model, including representing the heat transport from rainfall caused by the temperature difference between raindrops and land surface, and modifying the snow cover/melt curve to produce more realistic surface snow coverage and reduce wintertime low bias in temperature.

The WRF model output is acquired from the NCAR Research Data Archive ds612.0 [105]. We take the pressure, temperature, mixing ratio, height from the CTRL 3D subset, and surface topography, surface pressure, 2m temperature, and mixing ratio from the CTRL 2D subset.

2.3 Methods

2.3.1 CAPE calculation

All CAPE values shown in this work are calculated with SHARPpy (the Sounding and Hodograph Analysis and Research Program in Python) version 1.4.0a4, a widely used collection of sounding and hodograph analysis routines designed to provide free and consistent analysis tools for the atmospheric sciences community [`https://github.com/sharppy/SHARPpy`, 8]. SHARPpy is an extension of SHARP, which was first released in 1991 [63]. CAPE in the SHARPpy package is calculated following the definition of [86] in which temperature is automatically corrected to virtual temperature [40]. The required variables are vertical profiles of pressure, temperature, height, and dew point temperature. Wind speed and direction are optional and we do not include them. The package can produce the CAPE of parcels either at surface level (SBCAPE), at the "most unstable" level (MUCAPE), or using the averaged properties of "mixed layer" (MLCAPE). SHARPpy is the most commonly used package in the CAPE literature [e.g. 54, 74], which provides a comprehensive list of convective indices as output.

We evaluate CAPE for all summertime profiles corresponding to radiosonde soundings other than those with the following exclusion criteria: 1) no surface-level measurements (7 soundings or 0.004% of the total); 2) fewer than 20 vertical levels of observations (0.74% of soundings), or 3) excessive discrepancy of relative humidity between the surface and one level above, i.e. $RH_{sfc} - RH_{lev1} > 65\%$ (16 soundings or 0.008%). An excessive RH gradient implies unphysical mixing; the exact threshold is somewhat arbitrary but is chosen to exclude outliers where CAPE > 20,000 J/kg. In some cases, radiosonde profiles involve missing values in the height variable, even though temperature, pressure, and humidity are reported. In these cases, we interpolate height based on pressure using the SHARPpy "INTERP" function.

2.3.2 Testing sensitivity to vertical interpolation

In the analysis here we interpolate only where data are missing in radiosonde profiles, using the SHARPpy "INTERP" function. The number of vertical levels used is therefore inconsistent across datasets. Other authors of CAPE comparison studies have chosen to interpolate to produce consistent vertical sampling, for example, [54] who use 202 fixed levels (2 and 30 meters, followed by 75 m spacing from 75 m to 15 km). We test the robustness of derived CAPE to this interpolation by considering mean errors in profiles binned by number of levels, using observations from the year 2012. We find that mean absolute errors introduced by interpolation are over 25 J/kg for profiles with <10 levels, but fall to ~5 J/kg once the number of levels exceeds 60. However, since profiles with more missing levels are more likely to have low CAPE, the mean fractional error drops even more steeply (Figure S1): 14% for <10 levels; 7% for 10–20 levels; and 0.8% for >20 levels, justifying our choice of 20 as a cutoff. (Note that the *bias* introduced by interpolation across these profiles is smaller, only -0.3%. See [25] for similar conclusions.)

2.3.3 CAPE definitions

CAPE is the potential buoyancy of a parcel lifted to its level of free convection, but the parcel considered may be located at the surface (SBCAPE), at the most unstable vertical level (MUCAPE), or may be a hypothetical parcel initiated using the mean state of the mixed layer (MLCAPE). All are standard outputs of SHARPpy, with the lowest 100 hPa used to calculate MLCAPE. (See [16] for discussion of alternate choices). The appropriate CAPE definition differs according to the scientific question addressed. Some authors argue that MLCAPE is most appropriate for characterizing the average properties of the parcel being initiated by convection, and MUCAPE may best capture convective extremes [28, 16, 14], but we focus on SBCAPE in this work for several reasons. First, most prior CAPE comparison studies have used either only SBCABE [e.g. 55, 54], or all three definitions [25, 127]. SBCAPE is the most widely used in the climate community [107, 116], and several common CR-closure convective parametrizations use SBCAPE [e.g. 149, 143, 136]. Finally, using SBCAPE allows the most straightforward bias attribution, since it allows us to test the effect of errors in surface properties alone.

To understand the implications of the different definitions, we compare surface-based CAPE with that of the most unstable layer, MUCAPE, the maximum possible value for each profile (Figure 2.1). Because our focus is on incidences of very high CAPE, we are especially interested in whether different CAPE definitions lead to different understandings of the "high tail" (defined as incidents above 90th percentile CAPE). In all datasets, the higher the CAPE value, the more similar SBCAPE and MUCAPE become (Figure 2.1a). In conditions conducive to extreme weather (> 4000 J/kg), SBCAPE and MUCAPE are essentially identical in reanalyses and the WRF model output. Radiosondes show a slightly larger

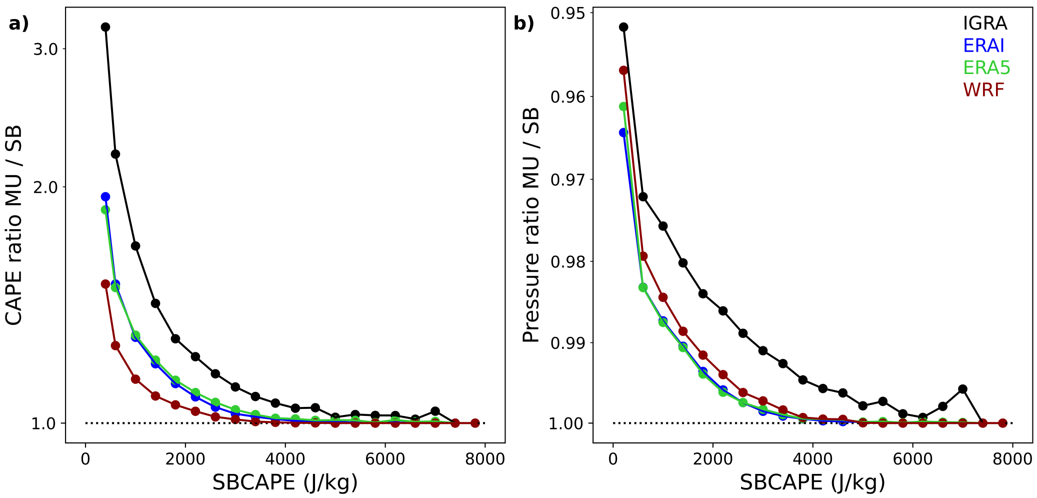

Figure 2.1: Comparison of SBCAPE and MUCAPE for all datasets, using all soundings considered. Data is binned by SBCAPE value, and we exclude values under 200 J/kg. **a)** Mean ratio of MUCAPE over SBCAPE, and **b)** mean of ratio of the most unstable pressure level over surface pressure. Note that y axes are log scale. For both CAPE and pressure level, the ratio approaches 1.0 as CAPE increases: in higher CAPE conditions, the most unstable level is closer to the surface.

distinction between SBCAPE and MUCAPE in all conditions. Mathematically, this means that model/radiosonde bias in MUCAPE must be more negative than those in SBCAPE. The pressure difference of the most unstable layer from the surface follows a similar pattern (Figure 2.1b). The higher the CAPE value, the more the most unstable layer approaches the surface, though observations again show more distinction. In conditions with SBCAPE ∼1000 J/kg, the average most unstable parcel in radiosonde soundings lies ∼30 hPa above the surface, but only ∼10 hPa in reanalyses and the WRF model. Above > 4000 J/kg, the most unstable layer in reanalyses and model lies at the surface.

In the last decade, some authors have argued that CAPE in any definition is not the best metric for diagnosing conditions conducive to severe weather. [59] propose using instead the maximum updraft velocity calculated with parcel theory, WMAX, which arguably better represents the intensity of updrafts. Several recent papers have followed that convention

[10, 102, 127]. Results here can translated to differences in WMAX distributions by using the approximation WMAX $\sim \sqrt{2CAPE}$, since the calculation of WMAX assumes no initial parcel vertical velocity.

2.4 Results – biases in CAPE distributions

2.4.1 CAPE distributions across datasets

Comparison of the distribution of CAPE in the datasets considered shows immediately that reanalyses and the WRF model output underpredict incidences of very high CAPE. Table 2.1 shows the breakdown of SBCAPE above or below threshold values, and Table 2.2 the same for MUCAPE. In all datasets, CAPE distributions are zero-peaked, i.e. a large fraction (~40%) of cases involve zero CAPE, even in the highly convective summertime. The frequency of zero CAPE is broadly similar across datasets, but in reanalyses and the WRF model, incidences of extreme CAPE drop off sharply, with values above 4000 J/kg substantially underpredicted in both definitions. For SBCAPE, reanalyses and the WRF model produce 40–50% fewer incidences of values $>$ 4000 J/kg. For MUCAPE, the underprediction is even more severe, with 65–75% of all incidences missed. (Biases in MLCAPE are intermediate between those in SBCAPE and MUCAPE.)

These biases in the "high tail" are related to a too-narrow distribution of CAPE in the WRF model and reanalyses. That is, reanalyses and the WRF model produce too few incidences of both extremely low and extremely high CAPE and too many incidences of intermediate CAPE. Figures 2.2 and 2.3 show distributions of non-zero CAPE values for SBCAPE and MUCAPE, respectively. Because valid zero values make up a large fraction of soundings, the choice whether to include them can potentially affect analysis, but in the datasets here, zero incidences are similar (Tables 2.1–2.2). We use two methods to show distributions: histograms (probability density functions, or PDFs) and quantile ratio plots.

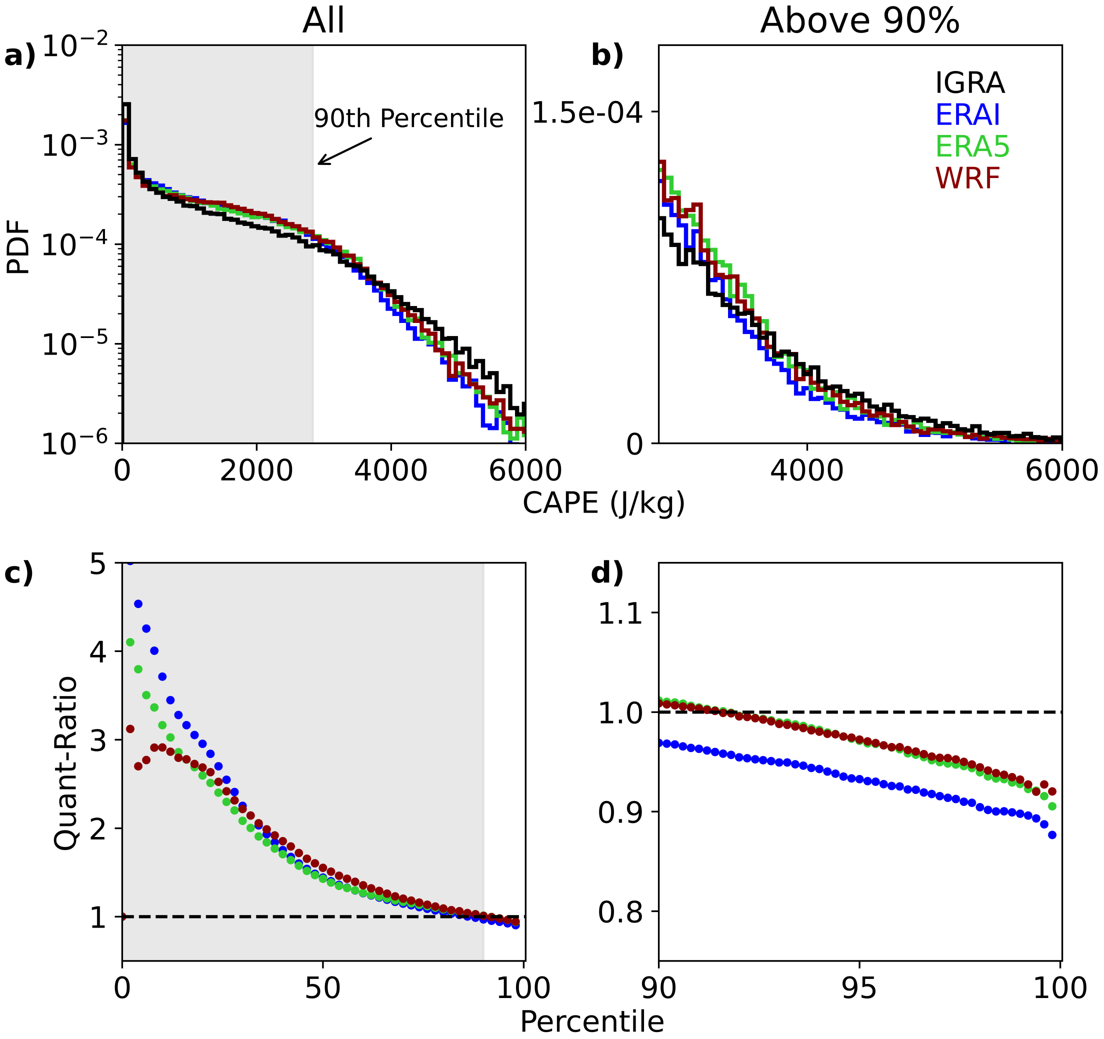

Figure 2.2: Probability density functions (top row) and quantile ratio plots (bottom row) of CAPE from reanalysis (ERAI and ERA5), high-resolution model output (WRF), and radiosonde observations (IGRA) for MJJA 2001-2012, with times and locations matched to IGRA observations. Points with zero CAPE are excluded (36-40% of datasets, see Table 2.1). Left column shows full distribution and right column the “high tail” (90th percentile and above). For IGRA, the 90th percentile is ∼2800 J/kg, the 95th ∼3200 J/kg, the 97.5th ∼4000 J/kg. In PDFs (top), plots are cut off at 6000 J/kg on the x-axis, omitting less than 0.1% of all points. In quantile ratio plots (bottom), a slope downward to the right indicates a narrower distribution. The WRF model and reanalyses consistently underpredict CAPE values in this “high tail”.

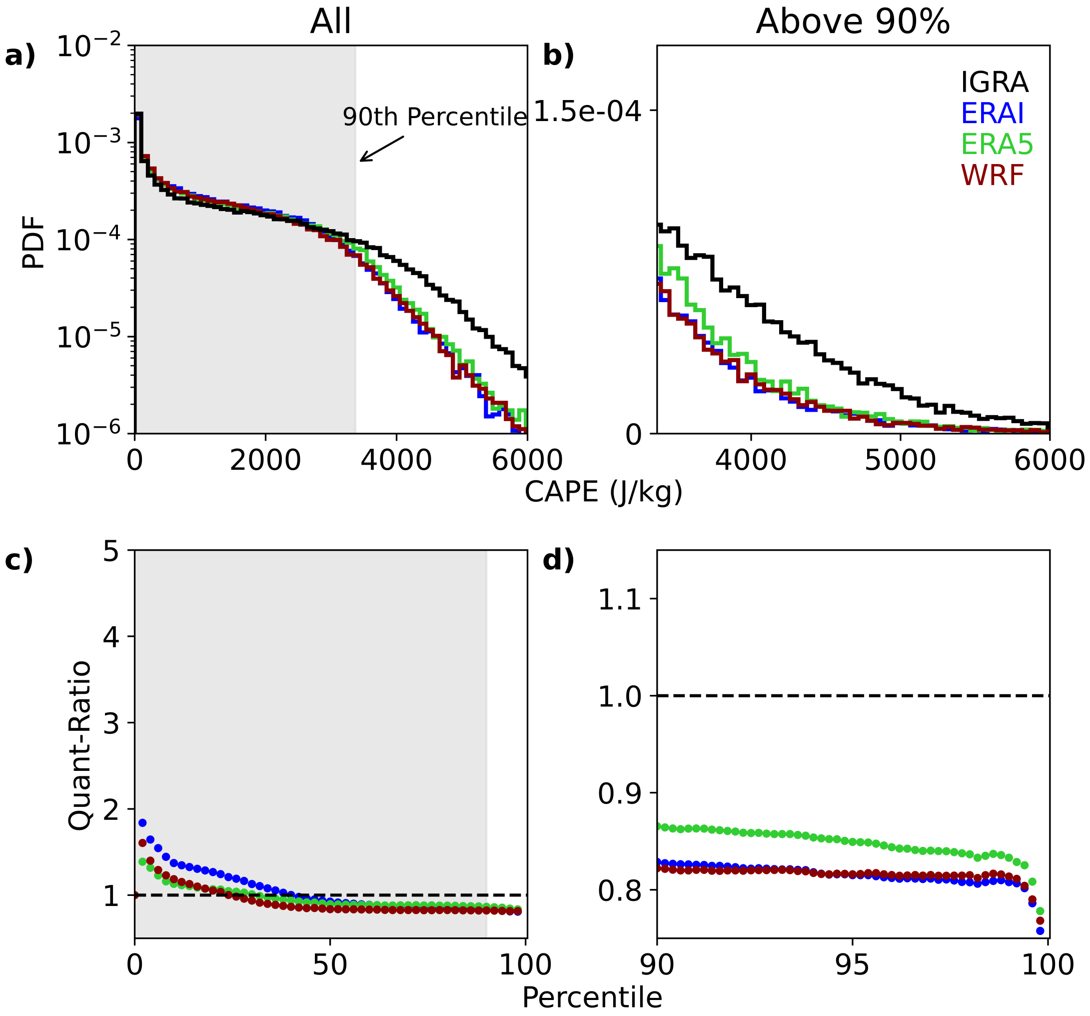

Figure 2.3: As in Figure 2.2, but for MUCAPE instead of SBCAPE. Points with zero CAPE are excluded from the analysis (23-35% of the datasets, see Table 2.2). We match the time and locations of model output to IGRA observations. PDF x-axes are cut off at 6000 J/kg, as less than 0.4% of all points lie above the limit. For IGRA, the 90th percentile is about 3370 J/kg, the 95th percentile ∼4010 J/kg, and the 97.5th percentile ∼4550 J/kg.

PDFs provide a basic sense of the CAPE distribution, and quantile ratio plots highlight distributional differences. Quantile ratio plots are constructed by taking the ratio of individual quantiles of two distributions being compared (e.g. CAPE in reanalysis and radiosondes); a value above 1 means that given quantile is overestimated. A simple multiplicative transformation produces a horizontal line whose value is the ratio of means, and a too-narrow distribution produces a slope downward to the right.

Table 2.1: Fraction of observations of SBCAPE in each dataset that exceed threshold values, or have zero value. Data used is the full 2001-2012 MJJA dataset, inclusive of zeroes, with time/location matched to radiosonde observations. Parentheses show the ratio of incidences observed for each model or reanalysis relative to IGRA radiosondes; a number smaller than 1 means underestimation. Note the large deficits in the most extreme SBCAPE category (>4000 J/kg), with the number of incidences underestimated by ∼40–50%.

	IGRA	ERAI	ERA5	WRF
Zeroes	36.1%	38.1%	35.0%	39.1%
> 2000 J/kg	13.0%	12.8% (0.98)	13.8% (1.06)	13.2% (1.02)
> 3000 J/kg	5.4%	4.0% (0.74)	4.9% (0.91)	4.6% (0.85)
> 4000 J/kg	1.6%	0.8% (0.50)	1.0% (0.63)	1.0% (0.63)

Table 2.2: As in Table 2.1 but here for MUCAPE. Deficits in the "high tail" are larger for MUCAPE than SBCAPE, as expected based on Figure 2.1. Parentheses show the ratio of incidences observed for each model or reanalysis relative to IGRA radiosondes. The number of incidences of MUCAPE above the conventional severe-weather threshold (2000 J/kg) is underestimated by ∼25–35% and that of extreme MUCAPE (> 4000 J/kg) by ∼65–75%.

	IGRA	ERAI	ERA5	WRF
Zeroes	22.7%	30.3%	28.2%	32.8%
> 2000 J/kg	22.3%	16.3% (0.73)	17.5% (0.78)	14.8% (0.66)
> 3000 J/kg	10.9%	5.2% (0.48)	6.5% (0.60)	5.0% (0.46)
> 4000 J/kg	3.9%	1.0% (0.26)	1.3% (0.33)	1.0% (0.26)

Reanalyses and the WRF model output considered here show the downward and rightward slope characteristic of too-narrow distributions: values are too large in low quantiles and too small in high quantiles. SBCAPE in the 20th–60th percentiles (50–1000J/kg) is overestimated by 84–94%, but above the 95th percentile is underestimated by 6–10%. These distributional errors occur even though mean SBCAPE values are similar in all datasets: within +1 to +6% with zeroes included, i.e. slightly *larger* in reanalyses and the WRF model than in radiosondes). This distinction highlights the need for distributional analysis, since even severe distributional biases may not be reflected in mean values (shown in

Appendix Table S1).

The too-narrow distributions in reanalyses and the WRF model are also seen when alternate definitions of CAPE are used. (Figures 2.3 and S2 show MUCAPE and MLCAPE, respectively). However, MUCAPE and MLCAPE also show significant low mean bias, with mean values -20 to -28% for MUCAPE and -16 to -22% for MLCAPE. These low biases lead to even stronger deficits in the "high tail", with quantiles above the 95th underestimated by ∼18–20% in MUCAPE and ∼15–17% in MLCAPE.

2.4.2 Spatiotemporal structure

Biases might be expected to show spatiotemporal structure, since CAPE is strongly linked to spatially complex fields of temperature and humidity. This relationship is illustrated in Figure 2.4, which shows a summertime snapshot of surface values from the WRF simulation (SBCAPE, temperature, and specific humidity), coincident with the radiosonde launch time at which CAPE values are typically highest (00 UTC, late afternoon or early evening in the contiguous U.S.). The time period shown is affected by a frontal system that brings high humidity to the Southeast and high temperatures to the Central U.S. (See Appendix Figure S3 for a weather map.) CAPE reaches extreme values only where both temperature and specific humidity are high, resulting in strong spatial gradients and a narrow band of extreme CAPE extending from southeastern Texas to northern Mississippi.

Two processes appear to drive the spatially correlated CAPE errors in Figure 2.4: large-scale patterns of model bias, and mismatches in the location of fronts or other weather features associated with strong gradients. The former is clearly evident in Figure 2.4. The WRF model is too warm and too dry in the Central U.S., coincident with and likely causing a large region of underestimated model CAPE. The warm-and-dry bias in this WRF simulation is extensively documented [83, 87]. Error in front location, on the other hand, likely produces overestimation of CAPE in stations in Tennessee and Alabama in Figure 2.4. Large-scale and

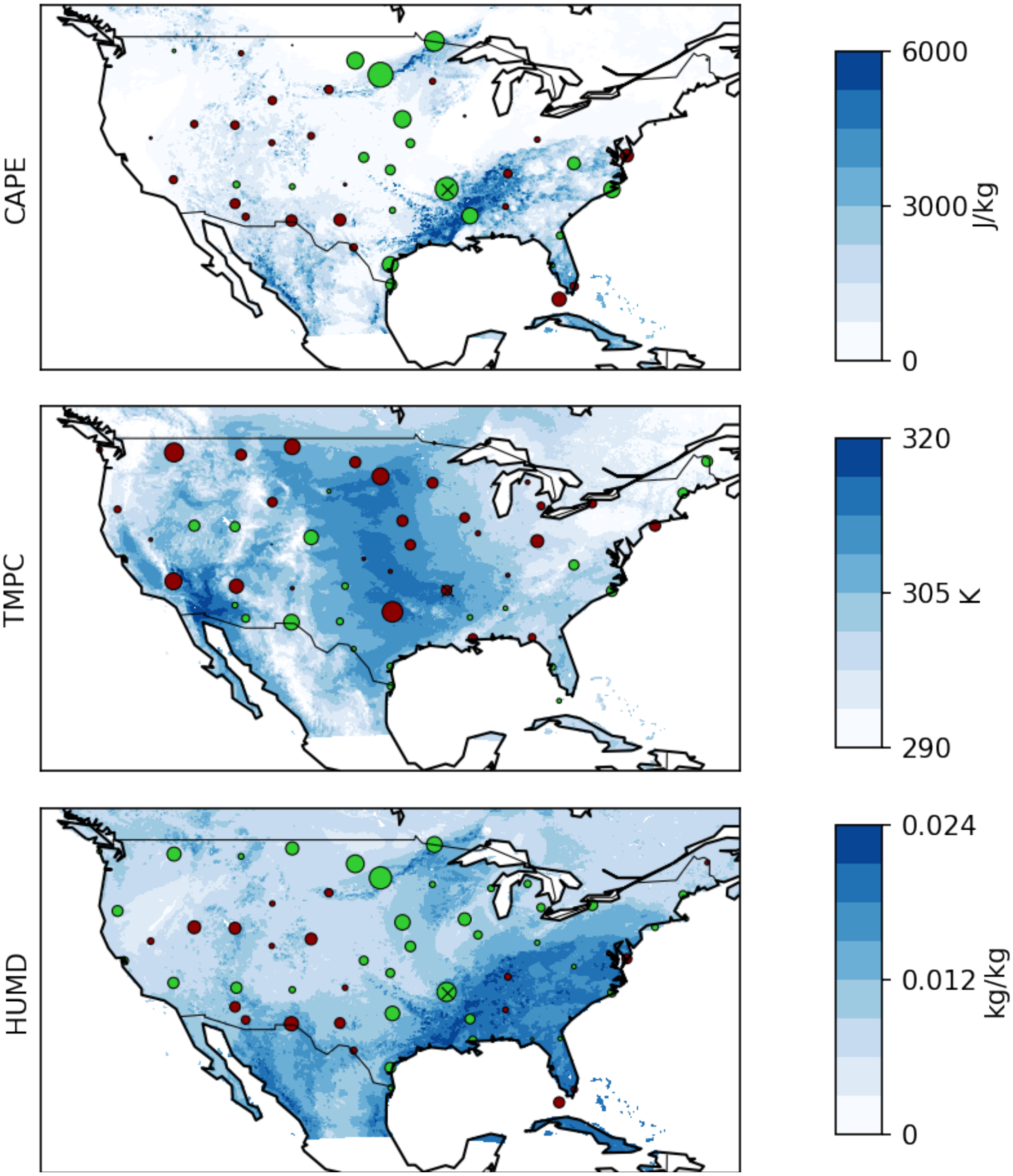

Figure 2.4: Snapshot of WRF simulation output at 00 UTC, July 21st, 2012. Panel colors show SBCAPE, 2 m temperature, and specific humidity. Ocean values are masked out. Circles show IGRA stations, with circle area showing the magnitude of bias in each variable and color indicating its sign (red = high, green = low). Note the low CAPE bias in the Central U.S. associated with too hot and too dry model conditions. Sounding marked “X” may also be affected by errors in the location of the warm front.

weather-related errors have different consequences for comparisons of CAPE in models and observations. Large-scale biases should be persistent, and will affect the overall distribution of CAPE. Fine-scale weather-related errors, on the other hand, vary rapidly on timescales of hours. While they can produce severe mismatch in individual soundings, and therefore introduce scatter in a model-observation comparison, they should have minimal effect on CAPE distributions.

2.4.3 Calibration with ground observations

Scatter in SBCAPE errors is in fact large in the WRF model and reanalysis products considered here, with correlation coefficients against radiosonde values of only R = 0.68–0.83. Figure 2.5 shows the comparison of WRF and radiosondes (panel a, R = 0.68); see Appendix Figures S4–S5 for ERAI and ERA5. Similar behavior is found in other studies, [e.g. 55] find correlation coefficients of 0.36–0.71; [127] find 0.71; and [54] show that reanalysis and satellite pseudo-soundings cannot reproduce radiosonde observed SBCAPE at individual timesteps.

Following [54], we test to see if these inaccuracies can be corrected by simply replacing surface thermodynamics fields with those from radiosondes (Figure 2.5). That is, we test whether errors in the WRF model and reanalysis SBCAPE are driven primarily by surface conditions rather than by the structure of atmospheric profiles. Both factors can be important because CAPE is a function of the integrated buoyancy across the convective layer, which is determined by both parcel and environmental temperature and moisture. In Figure 2.5, we successively replace surface values in WRF output, first temperature and pressure (panel b), then specific humidity and pressure (panel c), then all surface fields (panel d).

Surface values do seem to govern SBCAPE bias almost entirely. For WRF, correcting the surface specific humidity raises the correlation coefficient from 0.68 to 0.91, and replacing all surface fields raises it to 0.99, removing scatter almost entirely. While correcting temperature does not raise the correlation coefficient in WRF, and instead lowers it to 0.65, for other datasets the temperature correction also contributes positively; see Appendix Table S2. We also consider an alternate measure of correspondence, the percentage of points that fall

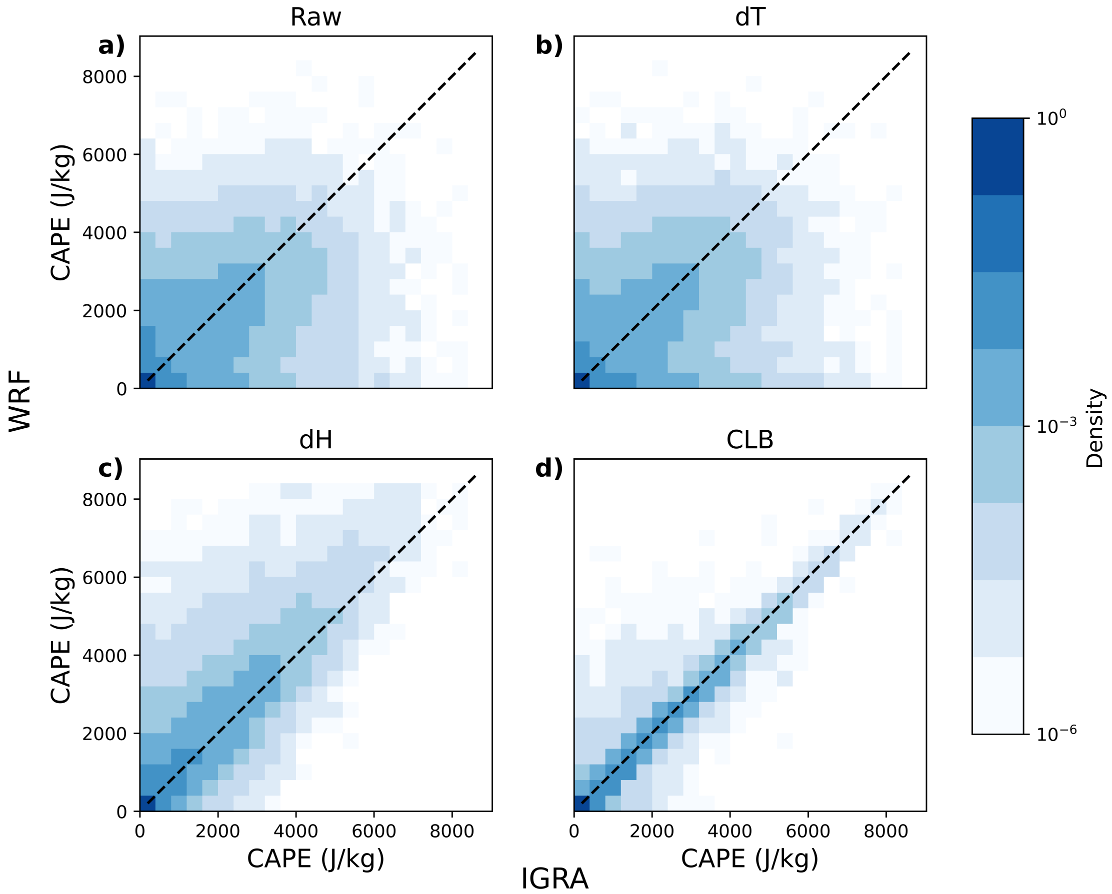

Figure 2.5: Comparison of SBCAPE in WRF and radiosonde observations, for all points during summer (MJJA) 2001-2012 when observations are available, inclusive of zeroes. Color bar shows log density (midpoint color is 1% of all observations), and both axes are also log scale. **a)** Raw data, showing wide scatter. Other panels: Recalculated WRF CAPE using **b)** observed surface temperature, **c)** observed surface humidity, and **d)** all surface values from observations. All recalculated CAPE values also involve a pressure correction whose effects are small. For analogous figures for ERAI and ERA5, see Appendix Figures S4–S5.

within ±800 J/kg of the one-to-one line (the width of two cells in Figure 2.5). For raw WRF data, the percentage is 78.6% (RMSE = 846 J/kg); correcting surface temperature raises the percentage slightly to 79.3% (RMSE = 875 J/kg); correcting surface humidity raises it to 90.2% (RMSE = 535 J/kg), and full calibration to 99.5% (RMSE = 162 J/kg). Results for ERAI and ERA5 are similar. Adjustment of surface values also largely corrects the distributional problems at high CAPE, so that for quantiles above 0.9, corrected SBCAPE values in reanalyses and the WRF model match those from radiosondes to within -0.2% to +2.0%. Correcting upper-tropospheric profiles has a minimal effect on CAPE values.

2.5 Results – CAPE in temperature & humidity space

The fact that reanalyses and modeled SBCAPE can be brought into agreement with radiosondes by simply replacing surface values implies that thermodynamic fields at upper levels are not important factors in SBCAPE biases. It may then be reasonable to consider SBCAPE as a function of surface thermodynamic fields alone. We therefore examine SBCAPE in the 2D parameter space of temperature (T) and specific humidity (H) to ask: 1) Is the density distribution of SBCAPE in T–H parameter space similar in reanalyses, model, and radiosondes? 2) What surface conditions are related to the highest SBCAPE days? and 3) What factors drive model and reanalysis biases in SBCAPE?

2.5.1 Dependence on surface temperature and humidity

CAPE distributions in T–H parameter space are in fact highly robust across all datasets. Figure 2.6 shows the heatmap of mean CAPE for radiosonde measurements, with data binned in steps of 3 K and 1.35 g/kg. CAPE values show a smooth gradient from lowest values at bottom left (warm and dry conditions) to highest at top right (hot and humid). This dependence on surface T and H is similar for all datasets (Appendix Figure S7). Contour lines at 2000 and 4000 J/kg for radiosonde observations are therefore nearly identical to those for other datasets (overlain). This similarity means that surface T and H robustly determine SBCAPE in all datasets. Of course, each T–H bin in Figure 2.6 involves an underlying CAPE distribution, but distributions are nearly identical for all datasets; see Figure 2.7 for two examples. These results support the previous finding that bias in SBCAPE can be explained by bias in surface measurements alone. (See Appendix Figure S8 for distributions of reanalyses and model errors in T, RH and H for the profiles shown in Figure 2.7, and

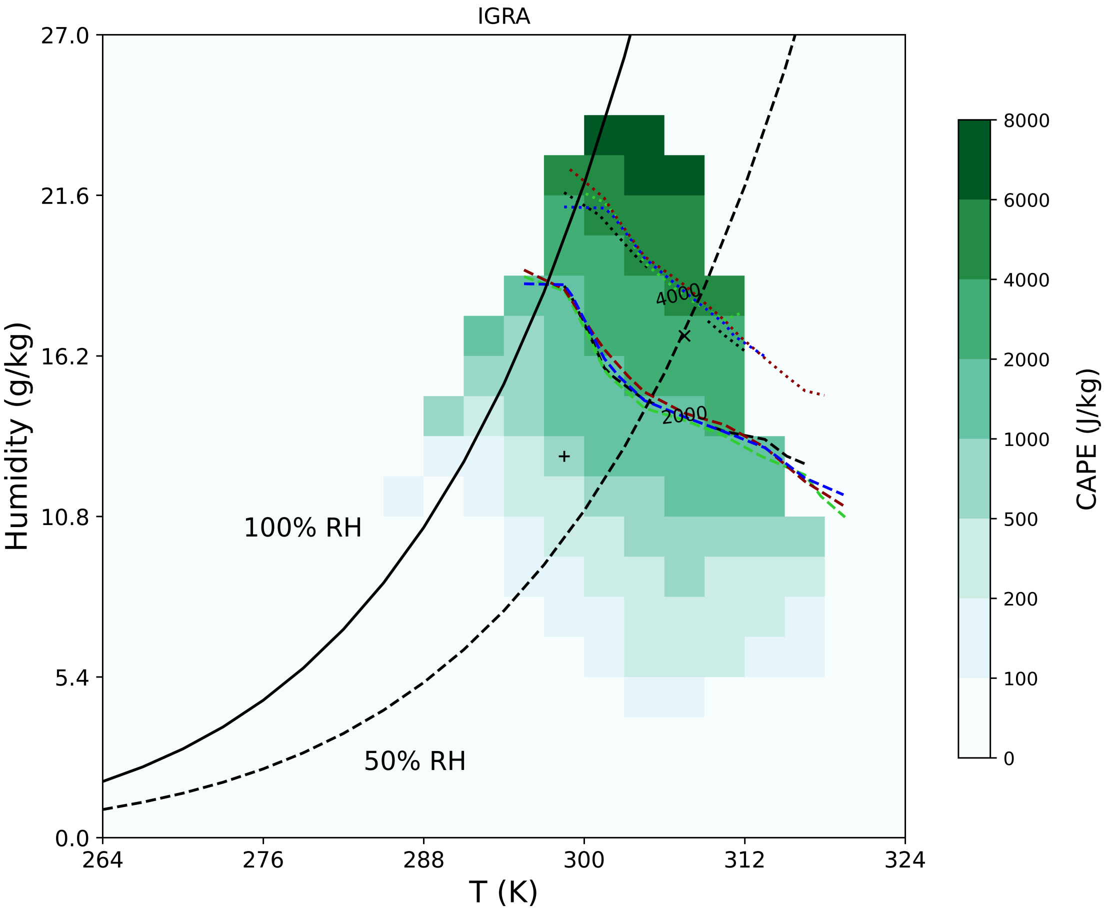

Figure 2.6: Mean radiosonde observed SBCAPE in surface temperature and specific humidity parameter space, for the entire dataset: summer (MJJA) 2001–2012 over the contiguous U.S., inclusive of all launch times and of zero values. Only bins with at least 10 samples are colored. Colors denote mean CAPE values averaged in bins of 3 K and 1.35 g/kg. Solid and dashed lines mark contours of 100% and 50% relative humidity (RH) at p = 1013 hPa. Soundings with lower surface p will be displaced up and left from these RH contours. Symbols '+' and 'x' mark two cases ('warm' and 'hot') used in Figure 2.7. Contours show approximate limits for 2000 and 4000 J/kg SBCAPE for all datasets with no surface corrections applied: IGRA (black), ERAI (blue), ERA5 (green), and WRF (red). Similarity of contours means that all datasets show similar bivariate distributions. See Appendix Figure S6 for the absolute occurrence in each bin in IGRA and Figure S7 for analogous figures for all datasets.

Table S3 for summary statistics.)

Only a restricted set of conditions tend to produce the "high tail" of CAPE distributions associated with extreme, high-impact weather. We show both 2000 J/kg and 4000 J/kg contours to bracket prior definitions of extreme weather thresholds. For example, [14, 134], and [35] all use 2000 J/kg in MLCAPE, which corresponds to SBCAPE ~3000 J/kg in our dataset. The conditions producing mean SBCAPE above 2000 J/kg involve temperatures above 297 K for 100% relative humidity (RH), or above 304 K for 50% RH. For mean SBCAPE above 4000 J/kg, the required temperatures are 2–3 K warmer, i.e. 299 K at 100% RH or 307 K at 50% RH. Significantly higher SBCAPE values are possible: in the most extreme conditions regularly sampled by radiosondes, 308 K at 65% RH, the average observed SBCAPE is over 7400 J/kg. Reanalyses and the WRF model rarely produce SBCAPE values this high (8 out of a million incidences, while observed incidences are nearly 10x more frequent at 60 out of a million), not because they differ in fundamental atmospheric physics but because they rarely sample the appropriate surface conditions.

2.5.2 Identifying sources of CAPE bias

Because SBCABE is strongly determined by surface temperature and humidity, biases in SBCAPE in reanalyses and the WRF model appear driven by biases in these surface thermodynamic values. We can therefore use the T–H diagram to identify the factors that lead to underprediction of the "high tail" of CAPE. Figures 2.8 and 2.9 use the same T–H diagram as in Figure 2.6, only now we show not the heatmap of CAPE but the density of observations of each T–H grid cell and the difference in that number between datasets. Because the diurnal cycle strongly affects surface values, we show separate figures for 00 UTC (U.S. late afternoon/evening) and 12 UTC (U.S. early morning), omitting the limited number of samples at other times. Reanalyses and the WRF model all underpredict the extreme T–H values associated with extreme CAPE.

Of the two times routinely sampled by radiosondes, the cooler 12 UTC launches – early morning in the U.S. – do not generally involve conditions associated with high CAPE (Figure 2.8a). Conditions at this time are almost never warm enough to produce SBCAPE >2000 J/kg, even though relative humidities are high, with a tight distribution centered

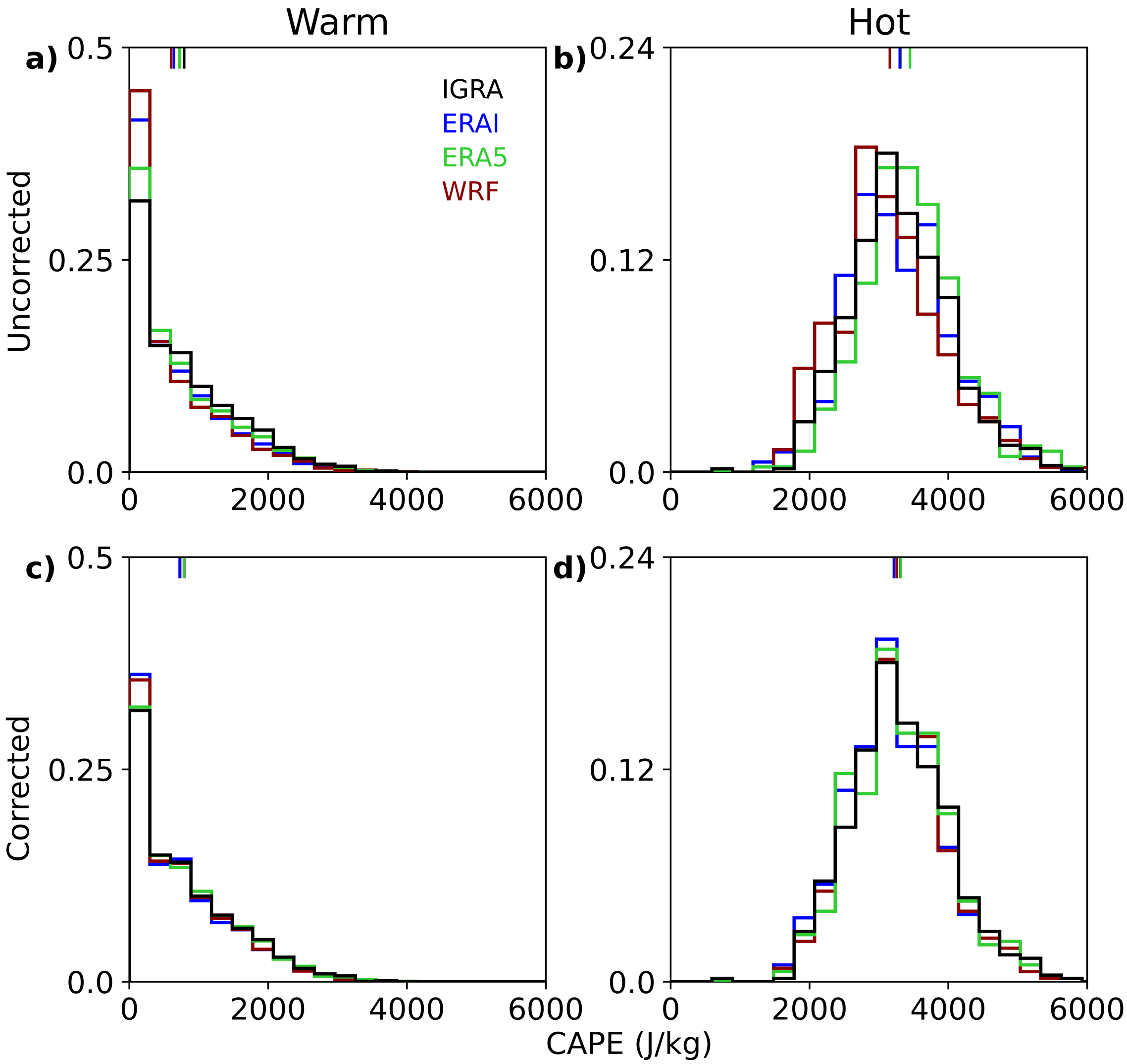

Figure 2.7: Comparison of SBCAPE in all datasets for specified T–H grid cell: the 'Warm' example is centered at 298.5 K and 12.825 g/kg (63.4% RH), and has mean SBCAPE 791 J/kg; the 'Hot' example is at 307.5 K and 16.875 g/kg (48.7% RH) with mean SBCAPE 3308 J/kg. Each bin is 3 K in width, and 1.35 g/kg in height. Top row shows uncorrected SBCAPE from reanalyses and the WRF model, and bottom corrected with IGRA surface values. Note that since the correction involves adjusting surface T and H, the profiles sampled in top and bottom rows are different. The 'warm' bin has 2438 profiles in the uncorrected data and 2063 in the corrected, while 'hot' has 378 and 508, respectively. Tickmarks at panel top show the mean of each distribution. Distributions are very similar; correcting surface values only slightly adjusts means (from a maximum bias of -5% in uncorrected data to -2% after correction).

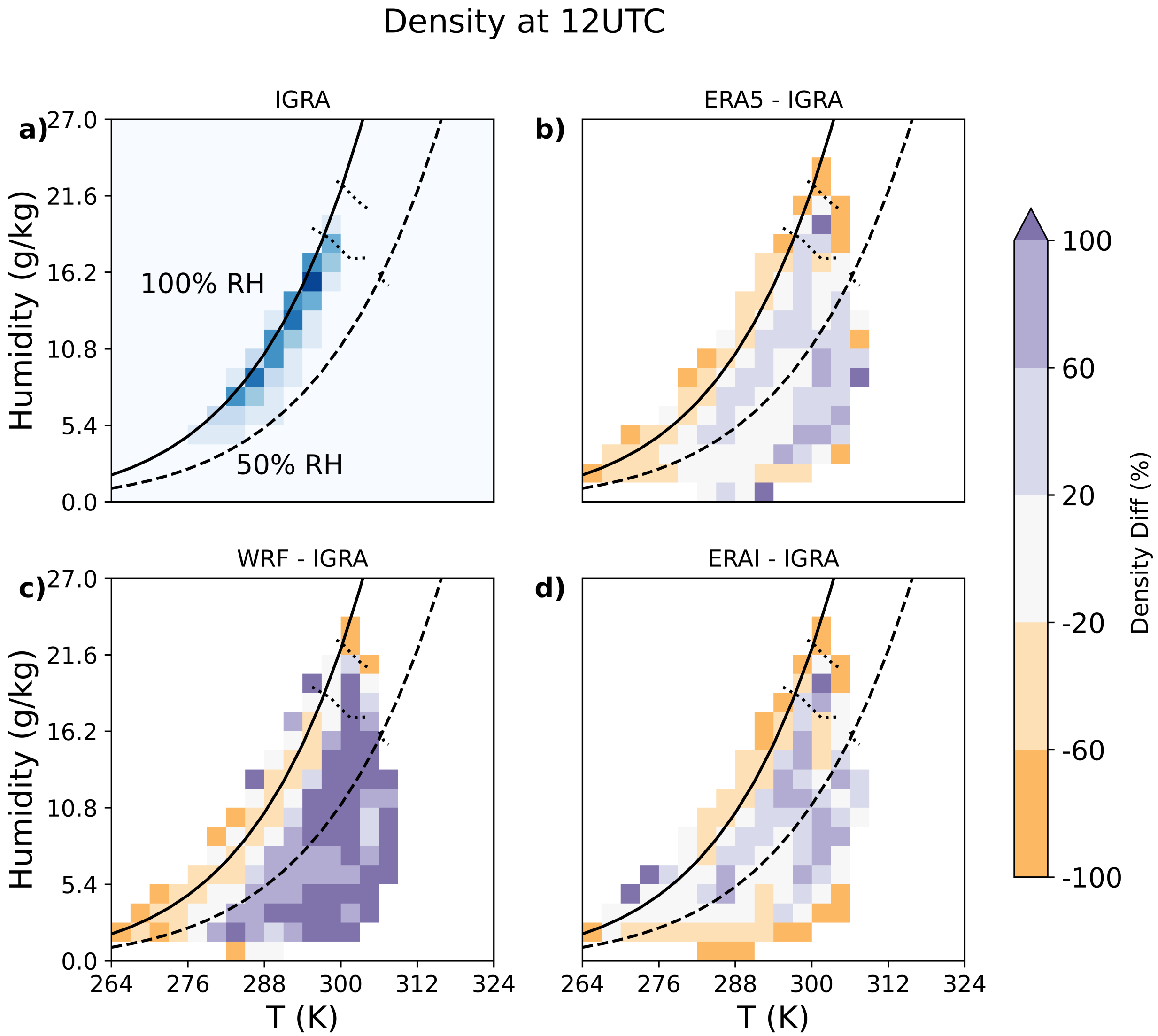

Figure 2.8: **a)** Density of observed surface conditions in temperature – specific humidity parameter space at 12 UTC (early morning in the contiguous U.S.), again for 2001-2012 MJJA radiosonde observations. Contours are repeated from Figure 2.6 to mark conditions associated with 2000 and 4000 J/kg SBCAPE. Darkest blue color shown is 5.6%–6.4% of distribution; lightest is 0–0.8%. Grids with no more than ten samples are defined as outliers and removed (only 0.03% of all model or reanalysis samples). Nighttime and early morning conditions are tightly distributed in relative humidity (RH ∼80%) and tend to be relatively cool ($T < 300$ K), with almost no conditions sampled that would tend to produce SBCAPE >2000 J/kg. **b–d)** – heatmaps of density differences between model/reanalysis and observations for ERA5, WRF, and ERAI. Color scale shows fractional difference after normalizing each bin by IGRA raw density. Orange = underpredicting and purple = overpredicting. Reanalyses and the WRF model all underestimate relative humidities (orange near the RH=100% contour) and WRF shows a strong warm dry bias (dark purple in lower right).

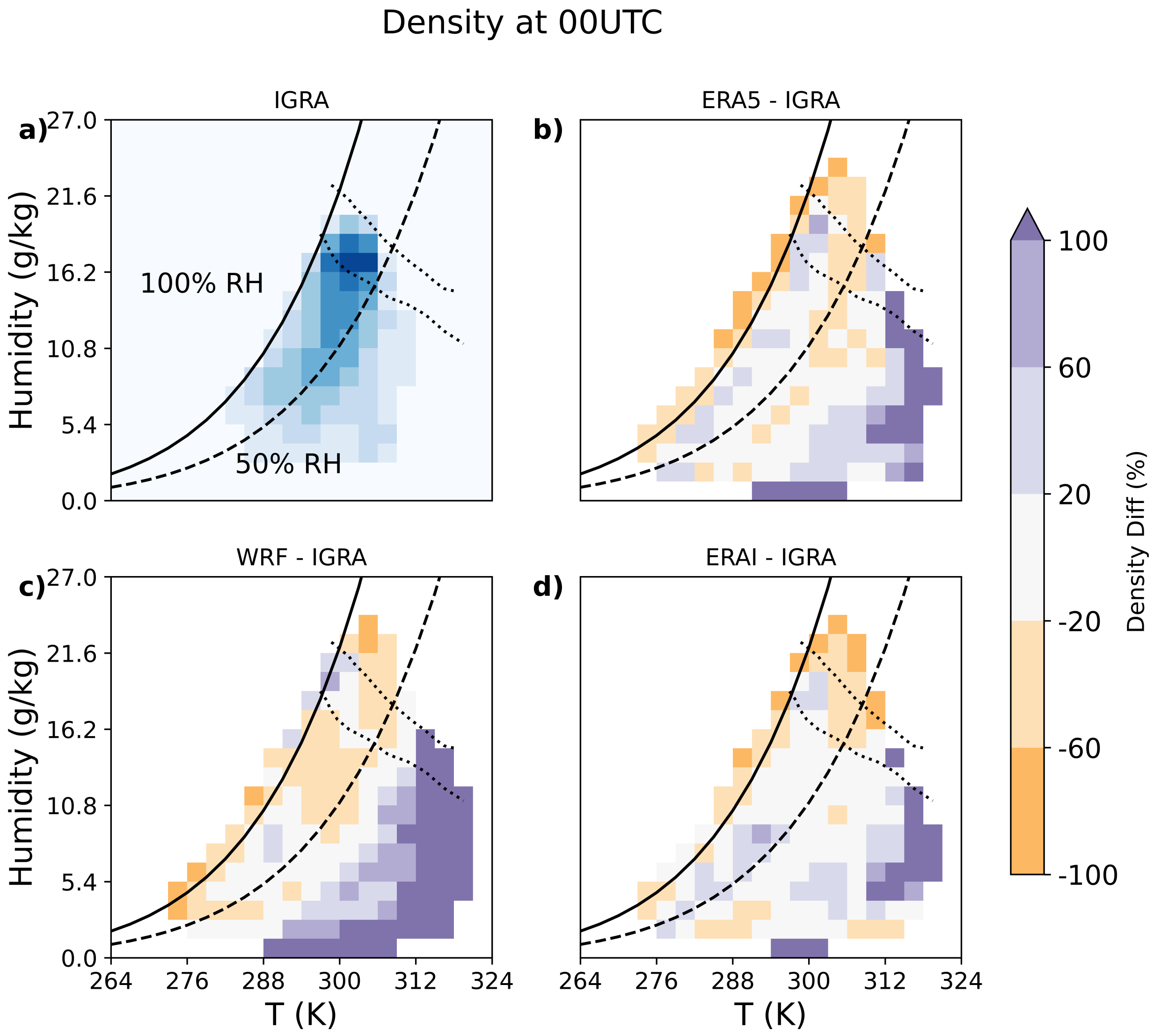

Figure 2.9: As in Figure 2.8, but for 00 UTC (late afternoon / early evening in the contiguous U.S.) **a)** Density of observed surface conditions in T–H diagram. At this time period the density distribution peaks in conditions associated with 2000–4000 J/kg CAPE. Darkest blue color shown is 2.1%–2.4% of distribution; lightest is 0–0.3%. **b–d)** Density differences between reanalyses / model and radiosondes. ERA5 and ERAI underpredict both the highest relative humidities and the highest temperatures (orange near the RH=100% contour and on the right side), while WRF shows a warm dry bias (purple in lower right). Reanalyses and model all severely underpredict the conditions associated with extreme CAPE (orange in upper right).

around ∼80%. Both reanalyses and WRF are biased dry, underpredicting incidences close to saturation, and WRF is also biased warm (Figure 2.8b–d).

Most of the observed extreme CAPE values occur during the warmer late-afternoon 00 UTC launches (Figure 2.9a). Relative humidities are lower then because specific humidity does not change much during daytime warming: the modal (most probable) 00 UTC surface conditions are between 303–309 K and ∼50% RH, with mean SBCAPE of ∼3000 J/kg, similar to the 'hot' example of Figure 2.7. Because reanalyses and WRF are dry biased, they underpredict high RH conditions in general and especially the extreme hot and humid conditions associated with the largest CAPE. On the other hand, reanalyses and WRF overpredict hot and dry conditions, WRF especially so (Figure 2.9b–d). The combined warm and dry bias explains why correcting WRF surface temperatures alone does not improve the match to radiosonde CAPE measurements.

2.6 Results – diurnal cycles of CAPE and biases

As shown in Section 5, the largest CAPE biases in the WRF model and reanalyses occur when conditions are most favorable to high CAPE, i.e. in daytime. This diurnal difference could result from inherent nonlinearity, but could also reflect a bias in some aspect of the diurnal cycle of surface thermodynamic fields. We therefore examine the diurnal cycle of surface temperature and specific humidity in reanalyses, model, and radiosondes. As an illustration, we show in Figure 2.10 a 5-day episode exhibiting large CAPE error, which is broadly representative of problematic reanalyses and model pseudo-soundings, and in Figure 2.11 we compare this episode to summertime climatological mean diurnal cycles for all, low-CAPE, and high-CAPE conditions (10th/90th SBCAPE percentiles). The sequence in Figure 2.10 runs from July 24–28, 2012 at a station in Wilmington, North Carolina. On 3 occasions radiosonde profiles show extreme CAPE of nearly 5000 J/kg, but model and reanalyses grossly underpredict these excursions, producing CAPE values ∼2500 J/kg too low. Since the temporal pattern of temperature evolution appears synchronous in all datasets, these biases appear unrelated to any mismatch of frontal systems. (See Figure S9 for the WRF model output and bias structure, and Figure S10 for the July 26th weather map.)

In the example episode of Figure 2.10, strong CAPE discrepancies result when models and reanalyses fail to capture short-term increases in specific humidity associated with extreme CAPE. Biases are driven by humidity, since throughout the 5-day period the WRF model and reanalyses are slightly too warm, with a fairly accurate diurnal temperature cycle (∼5-6K, with the WRF model exhibiting the largest amplitude). Comparison with climatological means on the T–H diagram of Figure 2.11 suggests that the daytime humidity rise in the example episode is extreme even for high-CAPE conditions, but the climatological biases are otherwise broadly similar. Reanalyses and WRF have overall dry biases that are exacerbated in WRF during the day (in the climatological case, by an actual daytime drop in specific humidity). Diurnal cycles of temperature cycles are similar, though daytime warming is slightly too weak in reanalyses, and WRF has an overall high temperature bias of ∼1.3 K. In all cases, too-low surface-level humidity appears to be the driving factor that strongly suppresses incidences of extreme CAPE.

CAPE biases in the example episode of Figure 2.10 differ in ERA reanalyses and the WRF model, but both produce deficits in specific humidity. In all datasets, temperatures match reasonably well in early morning (12 UTC), but daytime temperature rise is slightly too small in ERA reanalyses and considerably too large in WRF. ERA RH is reasonably accurate throughout, so its too-low temperatures are associated with a small specific humidity deficit. In WRF, specific humidity actually falls during the day, something not seen in reanalyses or radiosondes, contributing to erroneously low relative humidities. During the two "missed-high-CAPE" episodes, WRF RH is ∼25 percentage points below that in radiosondes.

To demonstrate that biases during this single-station episode are typical for warm condi-

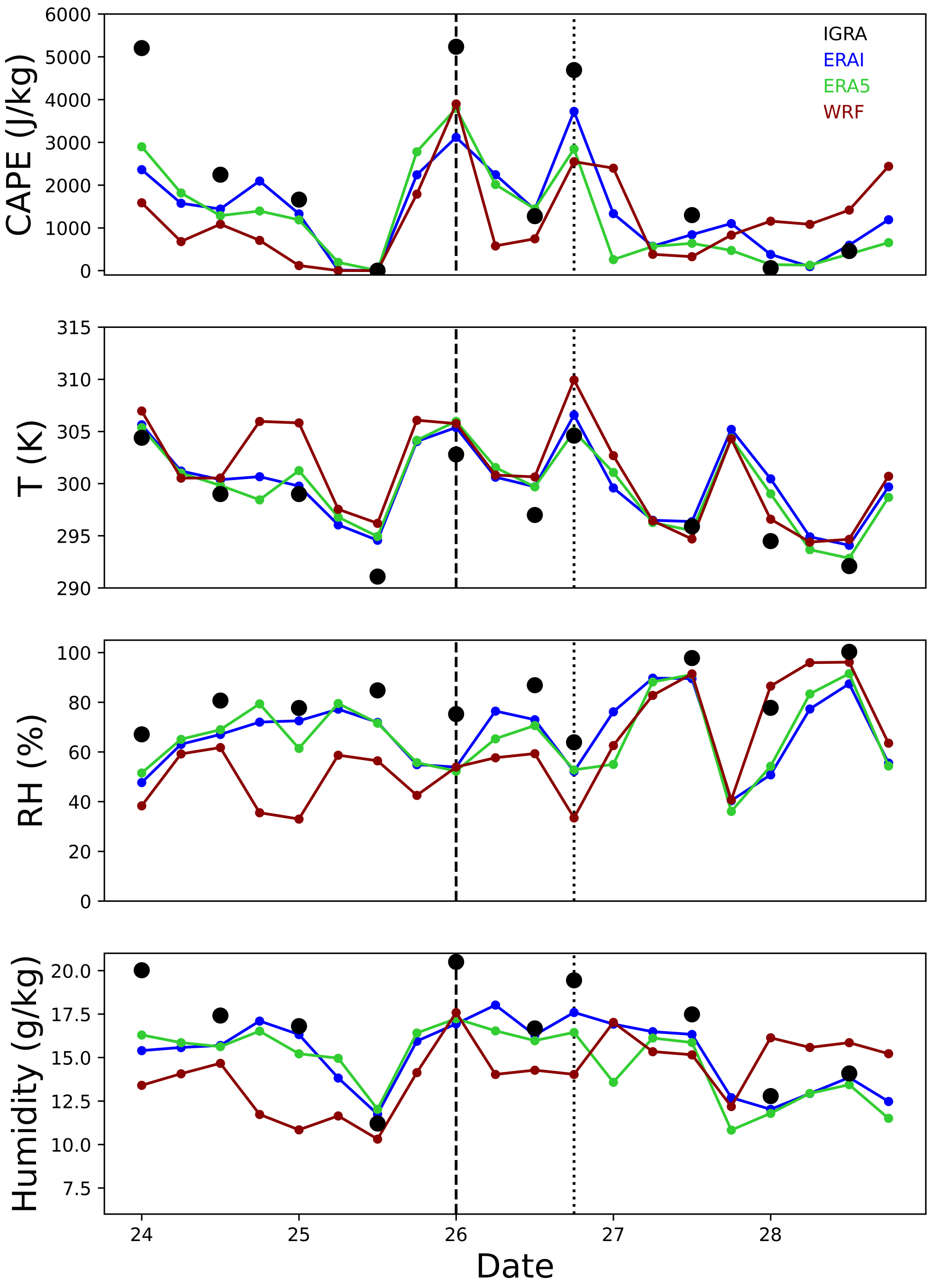

Figure 2.10: An example episode of high CAPE and substantial CAPE error: 5 days from July 24th to 28th, 2012 over Wilmington, North Carolina, color coded as before. Reanalyses and the WRF model are shown every 6 hours; IGRA soundings are generally every 12 hours (although note the irregular timing for radiosonde launching on the 26th, when daytime sampling occurred 6 hours earlier than usual, at 18 UTC). Vertical lines mark the two examples discussed in text. Over this entire period, reanalyses and WRF model show hot and dry bias; when the bias in absolute humidity is large, the dry bias produces too-low CAPE even despite too-high temperatures.

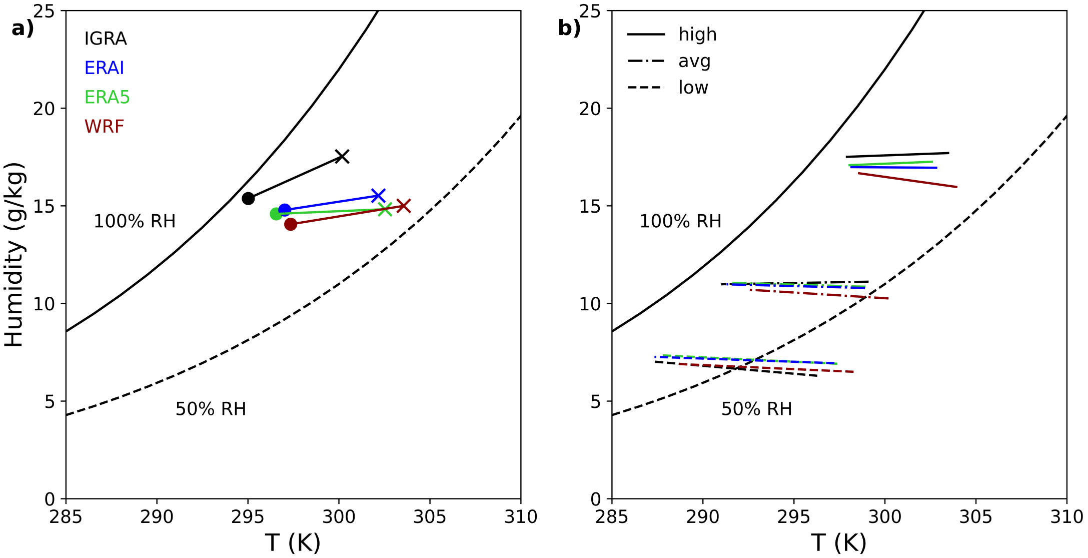

Figure 2.11: The diurnal cycle in T–H space in all datasets. Thick lines connect 00 UTC (right end, marked by 'x') and 12 UTC (left end, marked by 'o') values. Color code follows the convention throughout this work. **a)** The average over the 5-day episode in Figure 2.10. **b)** Mean summertime diurnal cycles over the entire domain, for all profiles (dot-dashed), and for high-CAPE (solid) and low-CAPE (dashed) subsets, defined as 00 UTC SBCAPE values above 90th / below 10th percentile in each dataset, and values 12 hours later. In all cases, WRF and ERA are biased dry, and WRF is biased warm. In the example (**a**), daytime observed specific humidity increases more than in WRF or ERA. In the climatological mean (**b**), specific humidity is roughly constant in observations and ERA but erroneously decreases in daytime in the WRF model, exacerbating mid-day bias.

tions conducive of extreme CAPE formation, we compare its diurnal cycle to climatological means across all stations. Figure 2.11 uses the temperature-humidity (T–H) diagram to show mean diurnal cycles of the 5-day episode at station Wilmington (panel a) and across the whole dataset (panel b), showing both the overall summertime average and also subsets of days involving the highest and lowest radiosonde SBCAPE values (90th/10th percentiles). The biases in the July 2012 example episode are similar to those generally experienced in high-CAPE conditions. On average, the ERA reanalyses experience slightly too-weak daytime warming (i.e. a too-small diurnal cycle in temperature) and are slightly too dry. WRF has an overall high bias in temperature of ∼1.3 K and is substantially too dry. The warm dry bias in WRF is exacerbated at mid-day by the fact that specific humidity erroneously drops during the day. This daytime humidity loss strongly suppresses incidences of extreme CAPE.

2.7 Conclusion and Discussion

Despite the importance of CAPE to both model construction and meteorology, few prior studies have evaluated CAPE biases against radiosondes on a large enough scale to evaluate climatological distributions. This study of nearly 200,000 proximity soundings in two reanalyses and a convection-permitting model confirms consistent patterns of distributional bias. CAPE distributions are too narrow in all cases, with underprediction of the most extreme values that are associated with severe weather events. Values in the 95th percentile and above are 6–10% too low in surface-based CAPE and even more severely underestimated under alternative definitions, at 18–20% too low in MUCAPE and 15-17% too low in MLCAPE.

In this study, both distributional biases and "mismatch error" in CAPE appear driven by conditions at the surface and/or boundary layer. SBCAPE shows a tight and similar dependence on surface temperature and humidity in all datasets; the dependence is so strong that CAPE distributions as a function of surface T,H are almost identical, even though individual profiles may have inaccurate surface values. The underprediction of the "high tail" of CAPE occurs simply because reanalyses and WRF runs underpredict the hot and humid conditions associated with extreme CAPE values.

These results emphasize the importance of land and boundary layer treatment in the performance of high-resolution models. Discussion of improving models has tended to focus on increased resolution and its effect on the atmospheric profile [e.g. 52]. However, the similarity of biases in model and reanalysis output with resolved and parametrized convection suggests that surface biases are unrelated to the treatment of convection. Many authors have noted that SBCAPE is strongly dependent on surface conditions [e.g. 84, 150, 37, 60], but land surface feedbacks may be crudely treated even in state-of-the-art high-resolution models [100].

Dry biases such as those seen here could be produced by misrepresentation of land-surface evaporation, by excess vertical mixing of the boundary layer, or, for the Central U.S., by too-weak advection of moisture from the Gulf of Mexico [50]. In this study, the greater bias in MUCAPE than SBCAPE across all datasets points to boundary layer processes as the problematic elements. It is well established that treatment of mixing in boundary layer (PBL) schemes can modify the diurnal cycle of temperature and humidity [73, 53, 144]. Several recent studies have evaluated the impacts of PBL treatment on CAPE, and consistently find that "local" schemes tend to undermix and moisten the boundary layer while "non-local" ones, like those used in the datasets shown here, overmix and dry it [26, 24, 48]. The YSU scheme used in our WRF runs is especially prone to producing a dry bias [26, 83].

While prior studies evaluating the effect of boundary layer treatment on CAPE have generally evaluated only mean values and have found only small biases, results here suggest that the "high tails" can be much more strongly affected. This finding is consistent with [48], who note in a small sample of soundings that dry biases produced by non-local

PBL schemes appear larger when observed CAPE is larger. The dependence of biases on underlying conditions means that even models and data products whose mean CAPE is well-validated may be inaccurate in capturing the strong convective events that lead to large socioeconomic losses. This problem cannot be assessed with studies that match soundings to severe weather events, since model displacement of weather features means that "mismatch" error is large and proximity soundings will not necessarily capture the same meteorological context. On the flip side, the accuracy of relatively low CAPE is also critical for convective parametrization schemes, since convective initiation thresholds are commonly set at only 65 J/kg. Subtle distributional biases can therefore affect convective triggering and total mass flux, and indirectly affect precipitation diurnal timing and amplitude. Given the importance of CAPE as a key meteorological parameter linking the large scale environment to weather-scale events, and its sensitivity to details of boundary layer treatment, its evaluation warrants careful distributional analysis.

2.8 Data Availability

The 4km high-resolution WRF simulations of the current and future climate of North America are available from NCAR Resaerch Data (`https://doi.org/10.5065/D6V40SXP`). The radiosonde observations are obtained from NOAA Integrated Global Radiosonde Archive (IGRA, `https://www1.ncdc.noaa.gov/pub/data/igra/`). The reanalysis output is available from ECMWF (`https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets`).

CHAPTER 3

ROBUST RELATIONSHIP BETWEEN MIDLATITUDES CAPE AND MOIST STATIC ENERGY SURPLUS IN PRESENT AND FUTURE SIMULATIONS

Published in *Geophysical Research Letters* in collaboration with:

Elisabeth Moyer[1,2]

[1]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA

Convective available potential energy (CAPE), a metric associated with severe weather, is expected to increase with warming, but we have lacked a framework that describes its changes in the populated midlatitudes. In the tropics, theory suggests mean CAPE should rise following the Clausius–Clapeyron (C–C) relationship at ∼6%/K. In the heterogeneous midlatitudes, where the mean change is less relevant, we show that CAPE changes are larger and can be well-described by a simple framework based on moist static energy (MSE) surplus, which is robust across climate states. This effect is highly general and holds across both high-resolution nudged regional simulations and free-running global climate models. The simplicity of this framework means that complex distributional changes in future CAPE can be well-captured by a simple scaling of present-day data using only three parameters.

3.1 Introduction

Convective Available Potential Energy (CAPE), loosely defined as the vertically integrated buoyancy of a near-surface air parcel, is a metric closely associated with extreme convective weather events that can cause substantial socioeconomic damages [e.g., 71]. CAPE is derived from the difference between the temperature profile of a parcel rising pseudo-adiabatically from the surface and that of the background environment [86], which determines the maximum possible updraft velocity during undiluted ascent. In meteorology, CAPE is used to predict thunderstorm events and in particular hail [58, 76, 72]. Studies have also used the covariate of CAPE and wind shear to explain differences in thunderstorm frequency across locations [14, 11] or across climate states [134, 35].

Early efforts to understand CAPE in observations sought to characterize it as a function of near-surface temperature and moisture [140, 147]. More recent studies of CAPE in observations have tended to focus on decadal-scale trends, often finding large increases. For example, [56] found trends equivalent to ∼50%/K in 15 tropical radiosonde stations. Model studies of CAPE under climate change have tended to produce smaller effects. Several recent studies that simulate the tropics using convection-permitting models (0.2–4 km resolution) without advection, i.e. approximating radiative-convective equilibrium, find CAPE increases of 8%/K [88], 8%/K [108], 12%/K [117], 7%/K [115], and 6–7%/K from theory [109]. In the midlatitudes, changes may be larger. [104] show 11%/K for 3 stations in the Eastern U.S.; [35] and [23] show ∼10%/K over the Eastern U.S.; and [80] find 10–14%/K for the entire U.S. These results are consistent across a wide range of model resolutions.

Theoretical frameworks to explain climatological CAPE fall into two groups. One approach assumes that background environmental profiles are fully determined by surface temperature, and predicts them by considering the effects of convective entrainment. [117] proposed a "zero-buoyancy model" based on the assumption that entrainment makes actual buoyancy in an ascending convective plume small relative to CAPE (with column RH considered fixed). [118] and [154] extended the work and validated the approach under radiative-convective equilibrium (RCE). However, the theory is not expected to work for midlatitudes land, which has strong spatial and temporal variations, even though its climatological mean profile is close to RCE [85].

A second approach treats surface and mid-tropospheric conditions as independent variables. [46] (henceforth EB96) drew on heat engine theory and described the relationship as

$$CAPE \;=\; A \cdot (h_s - h_m) \tag{3.1}$$

where h_s and h_m are moist static energy (MSE) near the surface (boundary layer) and in the mid-troposphere, respectively. In this perspective, CAPE represents the maximum possible kinetic energy that can be released given a heat transfer of $(h_s - h_m)$, and CAPE is generated only when surface MSE exceeds that of a mid-tropospheric threshold. [2], [82] (hereafter, AE17 and LC21) modified the approach to use a different threshold term, dry static energy, and showed that results captured aspects of CAPE variations in the midlatitudes.

We modify the framework based on [44] and use as the threshold term the minimum "saturation MSE" h_m^* in the mid-troposphere, the moist static energy a parcel would have if saturated:

$$CAPE = A \cdot (h_s - h_m^*) \tag{3.2}$$

We term the difference $h_s - h_m^*$ the "MSE surplus". The integral form of this expression can be derived from the definition of CAPE given the assumption that the effect of water vapor on buoyancy is negligible. (See Appendix Text S2.1 and Figure S11.) We then simplify to a linear dependence (as in e.g. AE17) by replacing the integral with a difference at a single location. This assumption is valid as long as the shape of the environmental temperature profile does not vary strongly with h_s and can be folded into the slope A. The rationale for h_m^* as the threshold term can also be expressed intuitively: CAPE depends only on temperature differences, and above the level of free convection, the rising parcel is saturated and conserves h^*, so its difference with the environment should be taken with a comparable quantity. [152] used h_m^* as a threshold for convective instability over summertime mid-latitude land, but Equation 3.2 has not yet been evaluated as a framework for CAPE.

A sufficiently general framework should explain not only average CAPE, or CAPE in the average profile, but its variations across space and time in the highly heterogeneous midlatitudes. This generality is required for any application to extreme weather, since only the "high tail" of CAPE is associated with the severe thunderstorms that produce large socioeconomic impacts. Although no prior work has addressed future changes in midlatitudes CAPE distributions, studies suggest they may shift in complex ways. For example, [23] show that spatial patterns of CAPE changes over North America differ from those of present-day CAPE.

In this work, we use observations and model simulations to evaluate how CAPE changes under CO_2-induced warming, and to test whether the relationship of Equation 3.2 captures these changes. That is, we ask whether it robustly applies to current and future CAPE distributions across climate states. Furthermore, we ask whether robustness means that complex distributional changes can be reproduced by as few as three parameters derived from regional means. Our goal is to quantify changes in CAPE distributions in the midlatitudes and to provide a simple framework that explains them.

3.2 Data and Methods

3.2.1 Model output

Most analysis here uses high-resolution model output: a paired set of present and future dynamically downscaled simulations over continental North America from the Weather Research and Forecasting model (WRF, version 3.4.1) run at 4 km resolution. Both runs are described in [83] and are acquired from NCAR RDA [105]. The present-day simulation (CTRL) uses ERA-Interim reanalysis for initial and boundary conditions and for a large-scale spectral nudging (scales >2000 km) applied to levels above the planetary boundary layer, to match planetary-scale weather patterns. Small-scale processes can still evolve freely. The future simulation is a pseudo-global-warming (PGW) scenario, treated identically but with reanalysis adjusted by a spatially- and temporally- varying offset derived from the CMIP5 multi-model mean projection under RCP8.5, to reflect large-scale changes under increased CO_2. These runs have been validated against observations [137] and used in studies of future CAPE changes [125, 104]. In this work, we use the years 2001–2012 and the equivalent future period.

To test whether results apply generally to a diverse set of free-running models, we use 11 CMIP6 models, selected based on the availability of the 6-hourly output needed for CAPE calculation. Model biases range from -60–+1700 J/kg, with the best performance (MPI-ESM1-2-LR) comparable to WRF, at ∼30 vs. 14 J/kg [137, 22]. We use pairs of historical (2005-2014) and ssp585 (2091-2100) simulations [49]. To allow comparison with observations, we subset all model output to 80 grid points that match International Global Radiosonde Archive (IGRA) weather stations in North America, as in [137]. For consistency, we calculate surface-based CAPE in all runs using the same python package. For 'paired' comparisons, we match each profile in CTRL/historical with its equivalent in PGW/ssp585. As in prior studies, most analyses here use only the summertime (MJJA or JJA), when convection is most active.

3.2.2 Methods: regressions and subsetting

All linear fits in this work are made using binned median data, to homogenize CAPE sampling. All fits are computed using orthogonal distance regression (ODR), which is most appropriate in conditions where errors in both dependent and independent variables matter.

See [114] for discussion of ODR. When fitting to estimate the fractional change in CAPE between climate states, we use the entire dataset, and we divide by the overall mean temperature change (4.65 K in WRF runs) when giving values in %/K. However, many comparisons focus on convective conditions and therefore involve a subset of the data. For regressions of CAPE against MSE surplus, we impose an absolute cut at CAPE >1000 J/kg. In other cases we compute values for profiles above the 73rd quantile in CAPE, which corresponds to CAPE >1000 J/kg in the WRF CTRL run. When constructing synthetic profiles, we apply a temperature offset derived from profiles with CAPE >73rd percentile in each climate state (3.92 K in WRF runs), to best capture the change in convective conditions.

3.2.3 Synthetic profiles

To help understand the minimal information needed to reproduce future CAPE changes, we construct three synthetic CAPE distributions based on the WRF CTRL profiles.

1. For *Clausius-Clapeyron* scaling, shown for illustrative purposes only, we simply multiply each CTRL CAPE value by 1.33 ($= e^{0.061 \cdot 4.65}$, where 6.1%/K is C–C for the mean temperature of high-CAPE profiles, 301.8 K). We neglect several factors whose systematic effects on CAPE would largely cancel: the projected rise in the Level of Neutral Buoyancy (LNB) (+0.6%/K); the reduction in surface RH (-0.4%/K), and treating profiles separately (-0.1%/K).

2. For the *constant offset* case, we add a fixed temperature offset of 3.92 K to each CTRL profile at each level from surface to 200 hPa (near the LNB in the mean CTRL profile), then linearly interpolate to zero change at 75 hPa. We show cases with and without a surface RH adjustment of -0.9%, the mean change for profiles with CAPE >73rd quantile.

3. For the *lapse rate adjustment* case, we modify the *constant offset* procedure to also

include a change in lapse rate $\Gamma = (T_s - T_{200})/z_{200}$. That is, we linearly interpolate between a warming of 3.92 K at the surface and a similarly-derived 4.94 K at 200 hPa. We also apply the -0.9% surface RH adjustment.

For context, we also show predictions of the SO13 theory under a 4.65 K temperature rise. We derive entrainment rate parameters of 0.67 and 0.68 for the WRF CTRL and PGW runs, and use true LNB values for each profile. ([117] used a fixed entrainment parameter of 0.75 and a fixed LNB temperature of 200 K.)

3.3 Results

3.3.1 Changes in CAPE distributions

We begin our analysis by asking: in midlatitudes model projections, how much and how does CAPE change with warming? In the WRF model runs, average summertime CAPE rises by 10% per degree of warming (a 61% increase, from 684 to 1103 J/kg with a mean surface temperature rise of 4.65 K). However, an alternate approach that emphasizes changes in higher-CAPE conditions may be more appropriate, and we use it throughout this work. We perform an orthogonal regression on the density distributions of paired profiles in present and future runs, which yields a clear shift upwards even though weather systems are not identical in the two runs and the scatter is therefore large (Figure 3.1, left). The slope yields a CAPE increase of 8.0%/K (45% total). With either method, the change is larger than in Clausius Clapeyron (6.1%/K) or in the SO13 theory developed for the tropics (6.0%/K), but smaller than would result from simply changing surface values while leaving atmospheric profiles unchanged (11.7%/K in the *constant offset* synthetic, which adds a single ΔT to all levels in all profiles; see Figure S12). Midlatitudes atmospheric lapse rates have therefore lessened slightly in the future simulation, as expected.

Distributional effects in future CAPE changes can be readily seen by comparing values for

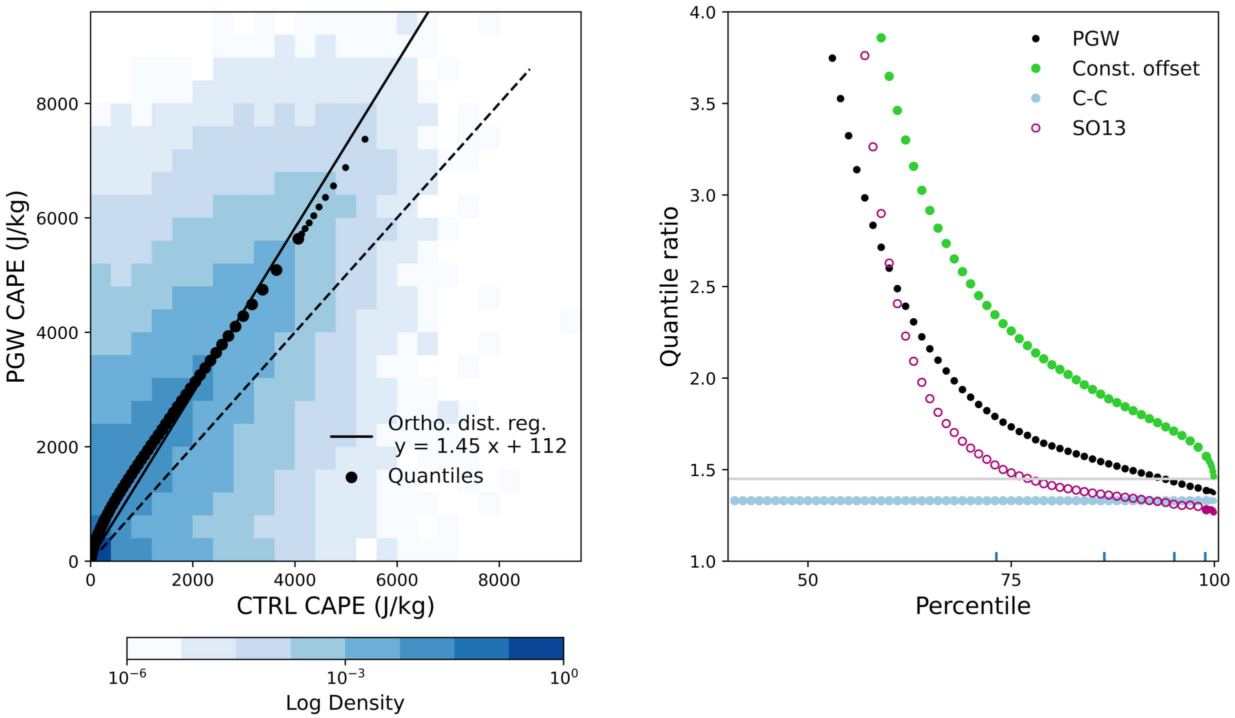

Figure 3.1: (Left) Comparison of CAPE in present (CTRL) and future (PGW) model runs as a density plot of paired profiles (see Methods), showing also the 1:1 line (dashed); the orthogonal regression (solid); and quantiles of the distribution (large dots, 1% increments from 0-0.99; small dots 0.1% increments above 0.99). (Right) Quantile ratio plot, constructed by taking the ratio of future to present CAPE quantiles, showing WRF output (black, same dots as L. panel), the synthetic datasets *C-C scaling* (light blue) and *constant offset* (green), and for reference *SO13* (purple, with changes computed relative to its own CTRL distribution). Gray horizontal line marks the +45% mean change from the orthogonal regression. Four vertical tick bars mark the percentiles matching 1000, 2000, 3000, and 4000 J/kg (73.2%, 86.5%, 95.1%, and 98.9%, respectively). The x-axis is truncated to omit quantiles where CTRL CAPE is zero. Changes in WRF are smaller than those in *constant offset*, implying some lapse rate adjustment.

individual quantiles to the overall regression line (Figure 3.1, left, dots). The lower quantiles lie above the regression line and the extreme high-CAPE quantiles (>∼3000 J/kg) below it, meaning the future CAPE distribution is narrower than that produced by a simple mean shift. This relative narrowing manifests as a downward slope in a quantile regression plot, which shows the ratio of individual quantiles of future vs. present-day CAPE (Figure 3.1, right). The effect is a necessary result of the nonlinear CAPE - temperature relationship: a given temperature rise produces a greater effect in low-CAPE conditions. For this reason, relative narrowing occurs even when surface temperature increases are uniform and environmental profiles do not change (*constant offset*, green) or in a theoretical approach that does not use observed environmental profiles (SO13, purple).

3.3.2 The effect of changes in environmental profiles

We found in section 3.1 that environmental adjustments appear to reduce future CAPE increases. To isolate this effect, we examine mean CAPE in surface temperature and humidity (T–H) space, following [137] (Figure 3.2). Since surface T and H uniquely define the moist adiabat on which a parcel rises, a change in CAPE for a given T–H is due only to an altered environmental profile. This approach effectively decomposes CAPE changes into a sampling effect and a partially compensating lapse rate effect. In the WRF model runs used here, increased sampling of hot and humid surface conditions in PGW would more than double CAPE from its CTRL values if environmental profiles remained constant (Figure 3.2, top), but environmental changes nearly halve that increase (Figure 3.2, bottom). This environmental damping makes future CAPE smaller for each T–H bin, so that hotter or wetter surface conditions are needed to achieve the same CAPE.

Most of this damping results from subtle changes in environmental profiles. Lapse rates across the domain lessen by 3% between CTRL and PGW, from -6.56 to -6.35 K/km (for the CAPE >73rd quantile subset). However, some damping also occurs even if the lapse rate distribution remains fixed (Figure S13). Because lapse rates in our domain are correlated with temperature – binned averages range from -5 K/km at 270 K to over -7 K/km at 320 K – then as the surface warms, each given temperature become associated with more stable conditions (Figure S14). The combined result is that CAPE contours in T–H space shift substantially between CTRL and PGW.

We can immediately make two inferences about CAPE changes in our model runs. First,

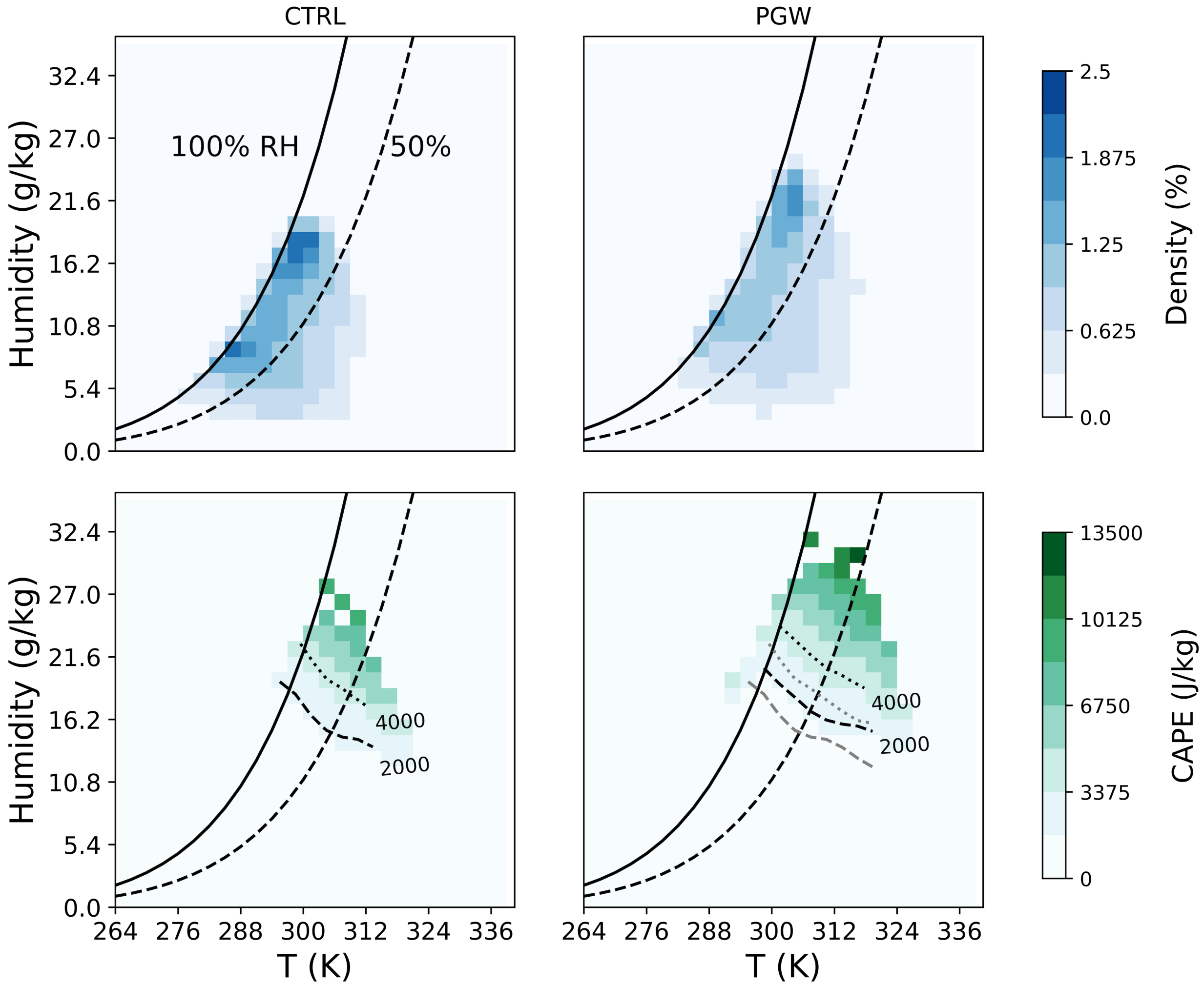

Figure 3.2: Density heatmaps of (top) sampling of T–H bins and (bottom) mean CAPE in each T–H bin, in CTRL (left) and PGW (right) WRF runs during summer (MJJA). Bins shown are all those with 3 or more observations. Solid and dashed lines mark RH of 100 and 50%. In the bottom row, dashed/dotted lines mark CAPE contours at 2000 and 4000 J/kg, with CTRL contours repeated in PGW panel as gray lines. Although conditions sampled in PGW are hotter than in CTRL (top), each given T,H bin is associated with smaller CAPE (bottom).

because CAPE contours align with those of MSE (Figure S15), CAPE in our dataset must be strongly related to surface MSE. Second, because CAPE contours in T–H space shift while MSE by definition cannot, this relationship must shift in future simulations. Both effects are consistent with Equation 3.2.

3.3.3 CAPE-MSE surplus framework

As predicted, the relationship between CAPE and surface MSE is reasonably linear in each climate state and shifts as the climate warms (Figure 3.3, top left). That is, CAPE on average does not develop unless surface MSE (h_s) exceeds some threshold, which changes between present and future simulations. This threshold, the x-intercept of the fitted regression, matches the mean minimum saturation MSE (h_m^*) in each climate state to within $< 0.3\%$. When CAPE is plotted against MSE surplus ($h_s - h_m^*$) instead, as in Equation 3.2, the relationship becomes robust across climate states and the residual variance becomes smaller, suggesting that this is a fundamental physical relationship (Figure 3.3, top right). On both measures, variance and robustness, the CAPE-MSE surplus relationship of Equation 3.2 outperforms the expression based on dry static energy as in [2] and [82] (Figure S16, which shows both WRF runs and observations). Fitted slopes are nearly identical in WRF CTRL and PGW runs and in observations (0.27 in all), and intercepts are nearly zero (0.7, 1.1, and 1.6 kJ/kg for CTRL, PGW, and observations, respectively). In this perspective, the effects of climate change reduce to a greater sampling of conditions with high MSE surplus.

The relationship described by Equation 3.2 applies across all models tested and appears remarkably robust not only across climate states but across locations and times. It holds in 11 free-running climate models from the CMIP6 archive (Figure 3.3, bottom), though they differ strongly in their CAPE distributions and projected changes: mean values over present-day summertime N. America range from 704 to over 2461 J/kg, and future changes range from 5-10%/K. Their CAPE-MSE surplus relationships also differ, with slopes of 0.22

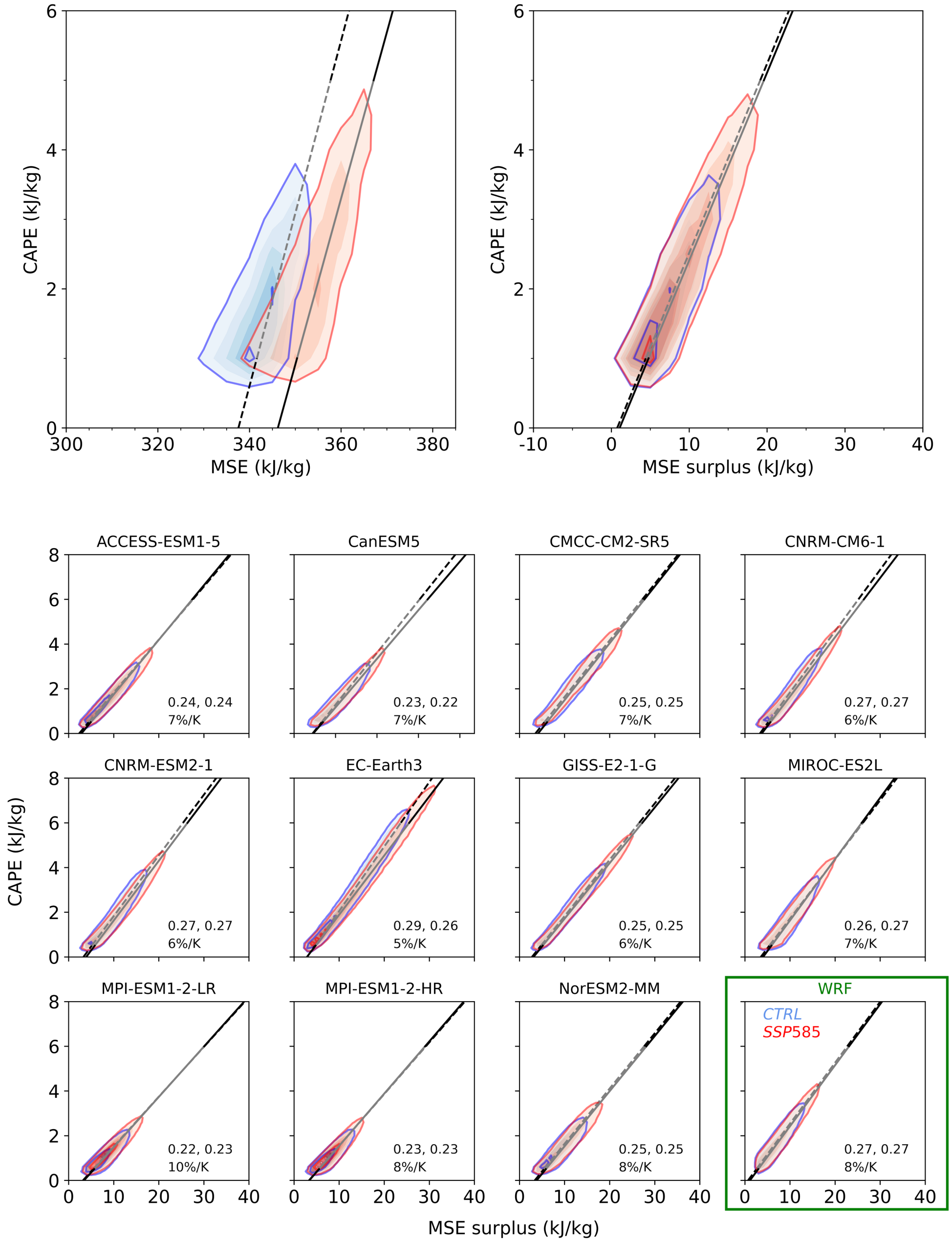

Figure 3.3: Cont. (Top) Relationships between CAPE and surface MSE (left) and MSE surplus (right), for WRF runs in N. America summertime (MJJA), showing all cases where CAPE >1000 J/kg (CTRL = blue, dotted; PGW = red, solid). Lines are fitted orthogonal regressions. Color shading increments are 1.5% for the left panel and 0.75% for the right. The CAPE-MSE surplus relationship is robust across climate states. (Bottom) CAPE–MSE surplus relationships in 11 free-running CMIP6 models and WRF for N. American summertime (JJA), using all cases where CAPE >500 J/kg. Color shading increments are 0.5% for all models except EC-Earth3 (0.25%). The CAPE–MSE surplus relationship is robust in all models, even those with with unrealistic CAPE.

to 0.29. Nevertheless, in each model that relationship remains constant across climate states. In the WRF model output, fitted slopes to CAPE vs. MSE surplus remain similar when the dataset is divided by latitude (northern vs. southern stations), by time of day (daytime vs. nighttime profiles), by interannual variations (anomalously warm vs. cold years), or even by season (winter vs. summer) (Figure S17).

3.3.4 A 3-parameter transformation

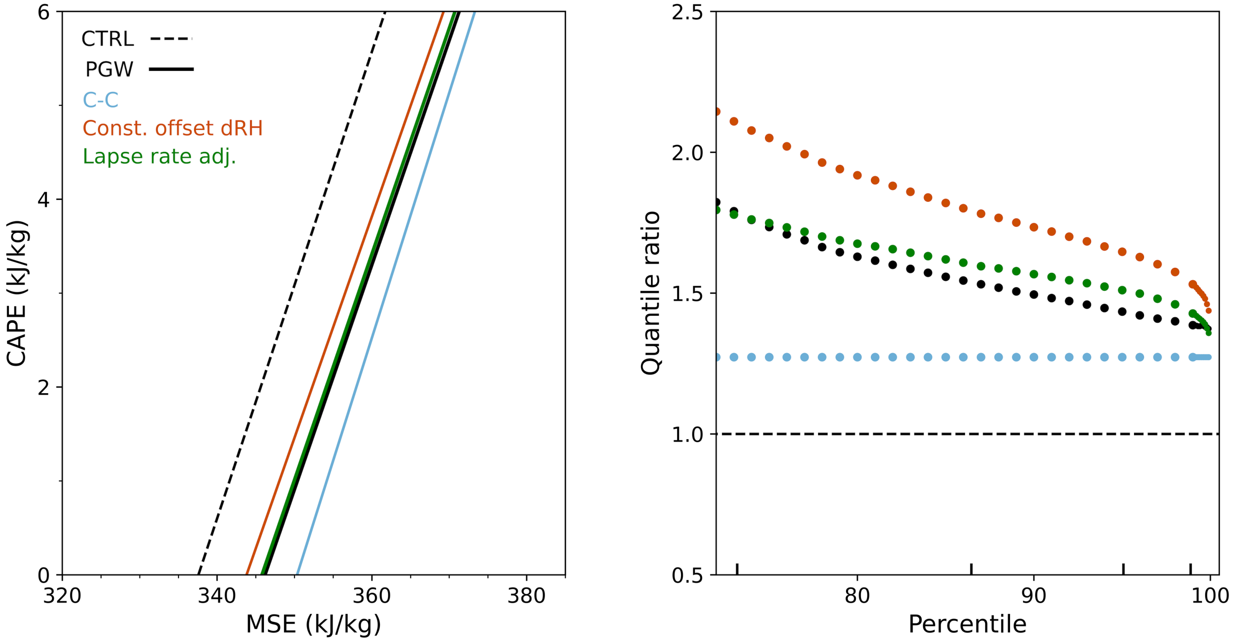

Figure 3.4: Comparison of present and future CAPE in model output (black) and synthetics: *C–C scaling* (light blue), *constant offset* including an RH adjustment (orange), and *lapse rate adjustment* (green). (Left) Fitted regression lines of the future CAPE-MSE relationship as in Figure 3.3. See Table S5 for slopes and x-intercepts. (Right) Future changes in CAPE as quantile ratio plots, as in Figure 3.1. The simple *lapse rate adjustment* effectively reproduces CAPE distributional changes.

The robustness of Equation 3.2 across climate states suggests that model-projected CAPE changes result from relatively simple adjustments. The fitted slope for each model, A, is a function of the shape of the environmental profile; for A to remain constant, that shape must not alter much. Changes in CAPE in Equation 3.2 can then result only from changes in surface conditions (h_s, which depends on surface temperature and humidity), or in a single metric of temperature in the free troposphere (h_m^*). While the quantile ratio plot in Figure 3.1 shows that transformations based on 1 or 2 parameters are insufficient for describing CAPE distributional changes, it appears that 3 parameters may be sufficient.

To construct our scaling, we use the two effects that produce the shift in CAPE contours in T–H space seen in Section 3.2 – an overall surface warming and a small decrease in mean lapse rates – and add the small but significant change in surface relative humidity in our WRF runs (-0.9%). As described in Methods, we calculate mean changes in these three parameters across our domain and apply them to the CTRL profiles. This simple adjustment correctly produces the shifting CAPE-MSE relationship, matching its slope and x-intercept (Figure 3.4, left). It also reproduces both the distributional narrowing and the magnitude of CAPE change for the high-CAPE conditions of interest (Figure 3.4, right). While midlatitudes CAPE is highly heterogeneous, a relatively straightforward transformation can capture its full distributional change in a future warmer climate.

3.4 Discussion

Increases in severe weather events, which are associated with high CAPE, are a substantial societal concern under global warming. Their understanding has been hindered by lack of a widely accepted theory or framework to describe midlatitudes CAPE changes. Theories developed for the convective tropics [e.g. 117] are not appropriate for midlatitudes land, where advection and a strong diurnal cycle mean that the mid-troposphere is often decoupled from the surface (Figure S19). In this work, we show that Equation 3.2, a modified version of the heat-engine theory originally proposed in 1996 (EB96) and of its later extensions (AE17, LC21), provides a compact representation of midlatitudes CAPE that is robust across space, over diurnal and seasonal cycles, and across climate states.

We term the work developed here a framework rather than a theory because the transformation requires empirical values and we do not predict the slope A, which accounts for the shape of the environmental profile and is empirically fit. Similarly, AE17 would require an empirical correction to their slope for a realistic moist atmosphere. In EB96, by contrast, A is based on thermodynamics and is effectively the Carnot efficiency of the atmosphere. In our WRF runs, the empirical slope of the CAPE-MSE relationship is larger than Carnot (0.24, vs. 0.14 for Carnot as defined by EB96), but this is not a violation of the 2nd Law given our focus on highly convective conditions.

Any transformation that describes changes in midlatitudes CAPE will necessarily require at least three parameters, one more than SO13 because the midlatitudes free troposphere cannot be predicted from surface T and RH even on average. In this work, we find that *only* three parameters are required: three regional mean values across our domain are sufficient to capture the full distributional change in the CAPE >73rd quantile. This result may seem counterintuitive since present-day North America encompasses a wide range of environmental conditions, future climate changes are spatially variable, and the response of CAPE is highly nonlinear. However, CAPE develops appreciably only in a relatively restricted subset of T–H space, where changes are more uniform.

The CAPE changes projected in our WRF runs and in most CMIP6 models are higher than Clausius-Clapeyron, the expectation under RCE. This difference matters for the occurrence of extreme conditions. Incidences of summertime CAPE >2000 J/kg, a commonly-used threshold for severe weather, rise half again as much in our WRF projections as under C–C scaling (14% in CTRL; >24% in PGW, 20% in *C–C*). Of course, predicting how this rise in extreme CAPE will affect future severe weather requires also understanding how it will map to a change in convective updraft velocities – but understanding CAPE changes under CO_2-induced warming is a necessary first step. The dependence of CAPE on MSE surplus provides a simple but robust framework for predicting and understanding that response.

3.5 Data Availability Statement

The 4-km WRF convection-permitting model output can be downloaded from NCAR RDA (`https://rda.ucar.edu/datasets/ds612.0/`). The IGRA radiosonde data can be downloaded from NOAA (`https://www.ncei.noaa.gov/products/weather-balloon/integrated-global-radiosonde-archive`). CMIP6 model output is available from the Earth System Grid Federation (ESGF, `https://esgf-node.llnl.gov/projects/cmip6/`).

CHAPTER 4

ROBUST CAPE SCALING ACROSS SPACE AND TIME

Prepared for publication in *Geophysical Research Letters* in collaboration with:
Daniel Chavas[3], Tiffany Shaw[1] and Elisabeth Moyer[1,2]
[1]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA [2]Center for Robust Decision-making on Climate and Energy Policy (RDCEP), University of Chicago, Chicago, IL, USA
[3]Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA

High values of Convective Available Potential Energy (CAPE) are often associated with severe weather, and the occurrence of such conditions is expected to increase with warming. Hence, it is important to understand how CAPE varies across space and different timescales (diurnal, seasonal, and climatological). In this work, we derive a CAPE scaling framework as a multiplicative combination of only a few physical parameters and show that the variations at a given location are primarily driven by convective layer depth and MSE surplus. We confirm the framework's validity with 6-hourly ERA5 and 11 CMIP6 models and show that the scaling reasonably captures full probability distributions, and spatial and temporal variations in the current climate. The framework also holds for the SSP585 future climate and predicts a 39% mean fractional increase in CAPE compared to a 42% modeled change averaged across 11 models. Changes are dominated by different parameters on different time scales: the MSE ratio dominates changes on climatological and diurnal time scales (>65%). On the contrary, the convective layer depth contributes to about 80% of the seasonal CAPE variations. The framework provides physical insights into how much and why CAPE changes across different time scales.

4.1 Introduction

High values of Convective Available Potential Energy (CAPE) are associated with a variety of severe weather events, e.g. severe thunderstorms, hail, and tornadoes [58, 76, 132]. Extreme convective events are associated with 15 out of 18 billion-dollar weather and climate disasters over the Contiguous United States in 2022 [120], resulting in huge socio-economic impacts across agricultural, transportation, and real estate sectors [139, 69]. To better understand the impact of convective extreme events, vast literature has attempted to understand the current climatology of CAPE [107, 4, 137] and how it relates to convective populations [14, 11].

Modeling studies have shown that CAPE is expected to increase as the climate warms [e.g. 80]. CAPE (in particular surface-based CAPE) is a strong function of surface temperature and humidity [38, 67, 115]; thus, its increase with warming can be largely explained by the fundamental physics of Clausius-Clapeyron scaling of atmospheric water vapor content, given that the relative humidity remains relatively unchanged as a constraint by the energy balance [64]. Furthermore, composite indicators for severe thunderstorm occurrence are largely driven by the increase in CAPE over the continental U.S. in century-long projections from CMIP5 ensemble [133, 35], and also shown to increase over Australia and Europe [4, 101]. However, few have attempted to attribute CAPE changes to different contributors. The closest attempt to our knowledge uses a set of convective parameters and state variables [23]. Even these attempts end up at a technical level and fail to provide a clear decomposition of actual CAPE changes across different scales. That is, a full scaling that works across all temporal and spatial scales of interest is needed to understand better factors contributing to CAPE changes.

Theoretical models for CAPE exist that provide insight into its variability, but all existing theories have some limitations. The simple theories based on zero-buoyancy model provides a useful prediction of CAPE for each climatological mean state and changes between climate states [117, 109], but the model ties mean temperature profile (thus CAPE) to entrainment and free-tropospheric moisture, which varies strongly across small spatial scales and on short (e.g., diurnal or synoptic) time scales. These assumptions are pertinent to all CAPE theories, and strong assumptions about water vapor are needed in this work as well. A recent attempt modeled the transient peaks over the mid-latitude continents, but its applicability to larger domains has not been tested [2, hereafter AE17]. [82, hereafter LC21] first derives CAPE scaling based on AE17 and confirms its applicability to spatial and diurnal variations for the continental U.S. However, LC21 is phrased in terms of dry static energy, which results in a less direct path to approximating CAPE than a framework based on moist static energy, as presented in this work. Additionally, LC21 requires using the entire free tropospheric static energy profile, whereas here, we seek to simplify the free troposphere further to use a single representative level.

[138] developed a scaling based on the strong dependence of CAPE on MSE surplus [$\Delta h = h_s - h^*_{min}$, see also 135] over Continental United States (CONUS) in summertime and finds this dependence holds across climate states, across a convection-permitting model and 11 CMIP6 models. This dependence is derived from Eq. (6.4.2) in [44], where CAPE is rewritten as a function of saturated entropy.The only necessary assumption is that the effect of water content on density is negligible. This assumption could lead to large errors for places where the environmental profile is drier and further from saturation but works reasonably well for convectively vigorous conditions. As in [2] and [138], we take out the MSE integral with a "shape of profile" parameter k:

$$CAPE \sim k\frac{\overline{\Gamma}}{\overline{T}}\Delta h \Delta z \tag{4.1}$$

where $\Delta z (= z_{LNB} - z_{LFC})$ is the convective layer depth, and $\Delta h (= h_s - h^*_{min})$ is MSE surplus (surface MSE subtracted by minimum environmental saturated MSE). $\overline{\Gamma}$ is mean environmental lapse rate; $\overline{T}$ is mass-weighted free-tropospheric temperature. Using an averaged $\overline{\Gamma}$ and $\overline{T}$ is sufficient for each evaluation (spatial or a temporal scale, see Table S6 and S7), suggesting that their variations have little contribution to the total CAPE variation at these scales [for Γ's effect, see 30]. The full derivation can be found in Text S3.1; an example profile is given in Figure S20. Note that the only free parameter is the "shape of profile" k, which partially compensates for the error in the assumption of negligible water vapor effect on density.

In this work, we will first test the robustness of the CAPE scaling (Eq. 4.1) in terms of full probability density function with ERA5 reanalyses and an array of CMIP6 models. We will then demonstrate the applicability of the scaling to CAPE variations across space, diurnal, seasonal, and climatological time scales. The simplicity of the scaling allows a linear decomposition of CAPE changes, serving as a tool to understand the main drivers of CAPE variations. The work aims to provide a robust and interpretable framework to understand better variations of CAPE across all spatial and temporal scales in current climate and across climate states.

4.2 Methods

4.2.1 Data description

We use General Circulation Models (GCM) runs from the 6th Coupled Model Intercomparison Project (CMIP6) [49] and ERA5 reanalysis data [65] for this work. The CMIP6 data used are 10 years of 6-hourly model-level output from historical (2005-2014) and ssp585 (2091-2100) simulations. Despite being a more extreme realization of future emission scenarios, ssp585 is used since it allows for the strongest forcing and is a stronger test for the theory [99]. The variables used to calculate the thermodynamic parameters (e.g., CAPE and MSE surplus) include temperature (ta) and specific humidity (hus). The manuscripts show results from JJA, and results for other seasons are shown in Supplemental Information. A detailed description of the simulations being used is included in Table S4, and the mean model biases for the full dataset are provided in Table S8. The ERA5 6-hourly output is acquired at 1°x 1° spatial resolution. ERA5 data matching the historical period used is used as a baseline validation of the representation of the current climate in CMIP6 models.

4.2.2 CAPE calculation

We use surface-based CAPE since past literature has extensively evaluated the mean and distributional biases in models against observation and reanalyses data products [e.g. 137, 22]. We follow these conventions and look at surface-based CAPE in this work, but the general methodology to construct such a framework should apply to mixed-layer and most-unstable CAPE. With a goal to calculate CAPE from a large volume of data in a fast and consistent way, we adapted the wrf-python (`https://github.com/NCAR/wrf-python`) package for this work [78], which originates from NCL and Fortran codes in WRF calculation. The parcels follow pseudo-adiabatic ascent, and virtual temperature corrections are implemented. The latent heat of freezing is not considered. We added functionalities called sbcape_2d and mucape_2d to the wrf module, and the source code is available from GitHub (`https://github.com/zwang02/wrf-python`).

4.2.3 Decomposition framework

Based on the CAPE scaling, we propose the decomposition of CAPE variations into two terms: the MSE surplus ($= h_{surp}$) and convective layer depth (Δz). Assuming the two terms are uncorrelated and that all other terms have little variation at the scales considered (represented by an error term), we have:

$$\frac{dCAPE}{CAPE} = \frac{d\Delta h}{\Delta h} + \frac{d\Delta z}{\Delta z} + \epsilon \tag{4.2}$$

Note that further decomposing the MSE ratio term without adding covariate terms is not practically useful or valid. The decomposition procedure is consistent with that in [22].

The decomposition assumes that variations in other terms, including column mean temperature ($\overline{T}$) and mean lapse rate (Γ), have a negligible contribution at the space and time scales evaluated here. For simplicity, we only use a constant for each model, each state that is considered in this work (e.g., MPI-ESM1-2-LR summer daytime in current climate $\overline{T} = 267$ K, $\Gamma = 6.61$ K/km).

4.3 Results

4.3.1 Applicability to full probability density function

We begin our analysis by validating the full probability distribution of CAPE (Figure 4.1), with 10-year (2005–2014) CAPE records in 11 CMIP6 models against ERA5 reanalysis (Methods). Four models (ACCESS-ESM1-5, CanESM5, EC-Earth3, and NorESM2-MM) perform reasonably well in capturing the distribution as in reanalysis data product (ERA5), with low mean bias below 50 J/kg. However, some models overestimate mean CAPE (highlighted by overestimating the "high tail" >2000 J/kg), with GISS-E2-1-G severely biased high (∼+480 J/kg relative to 400 J/kg mean CAPE in ERA5); in contrast, both MPI models miss the "high tail" and are biased low (both ∼-150 J/kg). We see diverse model performance in CAPE representation for the current climate (see Figure S21 for future climate), posing a challenge to reproduce CAPE characteristics faithfully.

Despite the challenge of authentic CAPE representation in CMIP6 models, the scaling (Eq. 4.1) captures the full distributions in each model and reanalysis product across the globe. In fractional measures, all models have a fractional mean absolute bias below 30%, with a few models (e.g., ACCESS-ESM1-5) performing very well in capturing the distribution as in reanalysis data product (ERA5) and their fractional bias falling below 20%. In absolute

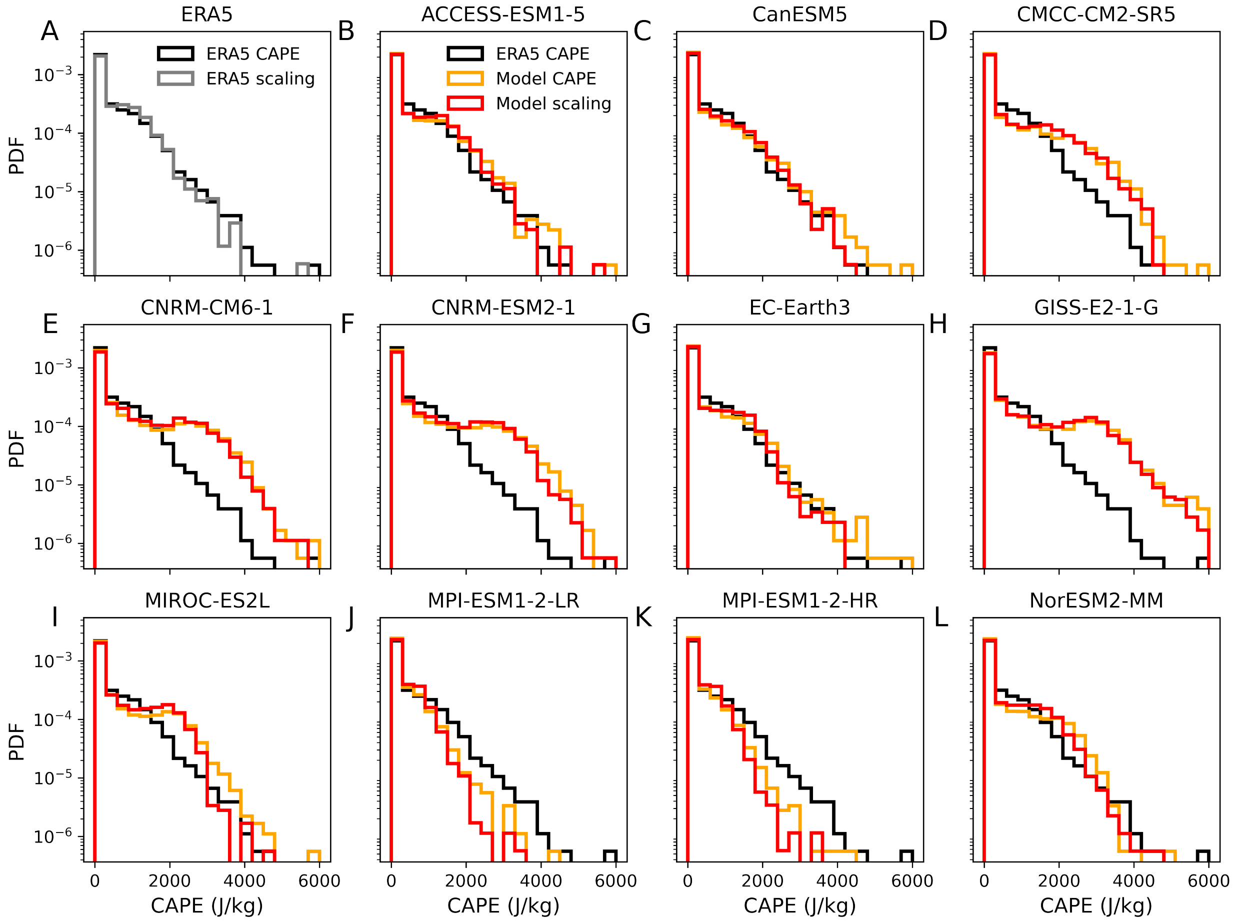

Figure 4.1: Probability distribution of CAPE from 6-hourly model data for the globe. We show ERA5 modeled (black) versus scaling (gray) in panel A and re-plot ERA5 as a baseline comparison in other panels. Each model uses a constant "shape of profile" parameter k fitted for the full distribution, with a spread from 0.52 to 0.61. We show the modeled CAPE (red) versus scaling (orange) of each CMIP6 model in all 11 CMIP6 models used in these panels and resample the dataset down to 6000 for each model and each variable. We use a semi-logy axis to highlight differences in the "high tail". The scaling consistently performs well, yielding a much smaller bias against the modeled value than the model bias against reanalysis (ERA5).

measures, we obtain mean absolute bias <100 J/kg for 6 out of 11 models with a constant k fitted for each model (see Text S3.2 for regression methodology). The performance is even better for the Contiguous U.S., with fractional mean absolute bias below 20% for all models (see Figure S22); the scaling still works reasonably by arbitrarily choosing a constant k = 0.5 for all models (see Figure S23). We also measure distributional differences with the 2-sample K–S test statistics (see Text S3.3 for details). 10 out of 11 models have a smaller D between scaling and modeled CAPE (at 0.078 when averaged across models) than that between models and reanalysis (at 0.183), suggesting our scaling can obtain a more faithful representation of CAPE given modeled conditions regardless of whether the model fails to represent the reanalysis authentically. The ability to reproduce modeled probability distribution suggests the scaling captures fundamental physics governing CAPE variations.

4.3.2 Applicability across space

A valid scaling should apply across space in a given climate state. Here, we evaluate the scaling against reanalysis or modeled spatial distribution of CAPE under the current climate state. Since CAPE is a strong function of near-surface conditions, it naturally exhibits strong spatial variation that, in turn, determines the occurrence and intensity of storms. Several regions across the globe are known to have vigorous convection throughout the year, e.g., the tropical ocean and maritime islands; conversely, multiple regions over the mid-latitudes land show strong seasonality, with average CAPE exceeding 500 J/kg in the summertime, including the Great Plains, the Amazonian, and Western Africa [107, 141, 75]. We evaluate the scaling with the correlation between scaling and modeled CAPE and mean bias for global and subregions of interest.

We first show that our regional scaling captures the spatial variation of the ensemble mean CAPE well (Figure 4.2, left). The need for regional scaling arises given the spatial heterogeneity of the "shape of profile" due to the dryness over land, we expect k to have a strong land-sea contrast but should be largely constant over the tropical ocean. That is, aimed to capture variations across land and ocean, the spatial extent to which assuming a constant k is acceptable needs to be determined. In this work, we show that k derived from a 21°x 21° box (see Text S3.4) sufficiently captures most of the spatial variability of CAPE (see performance of scaling and corresponding k with different assumptions in Figure S24).

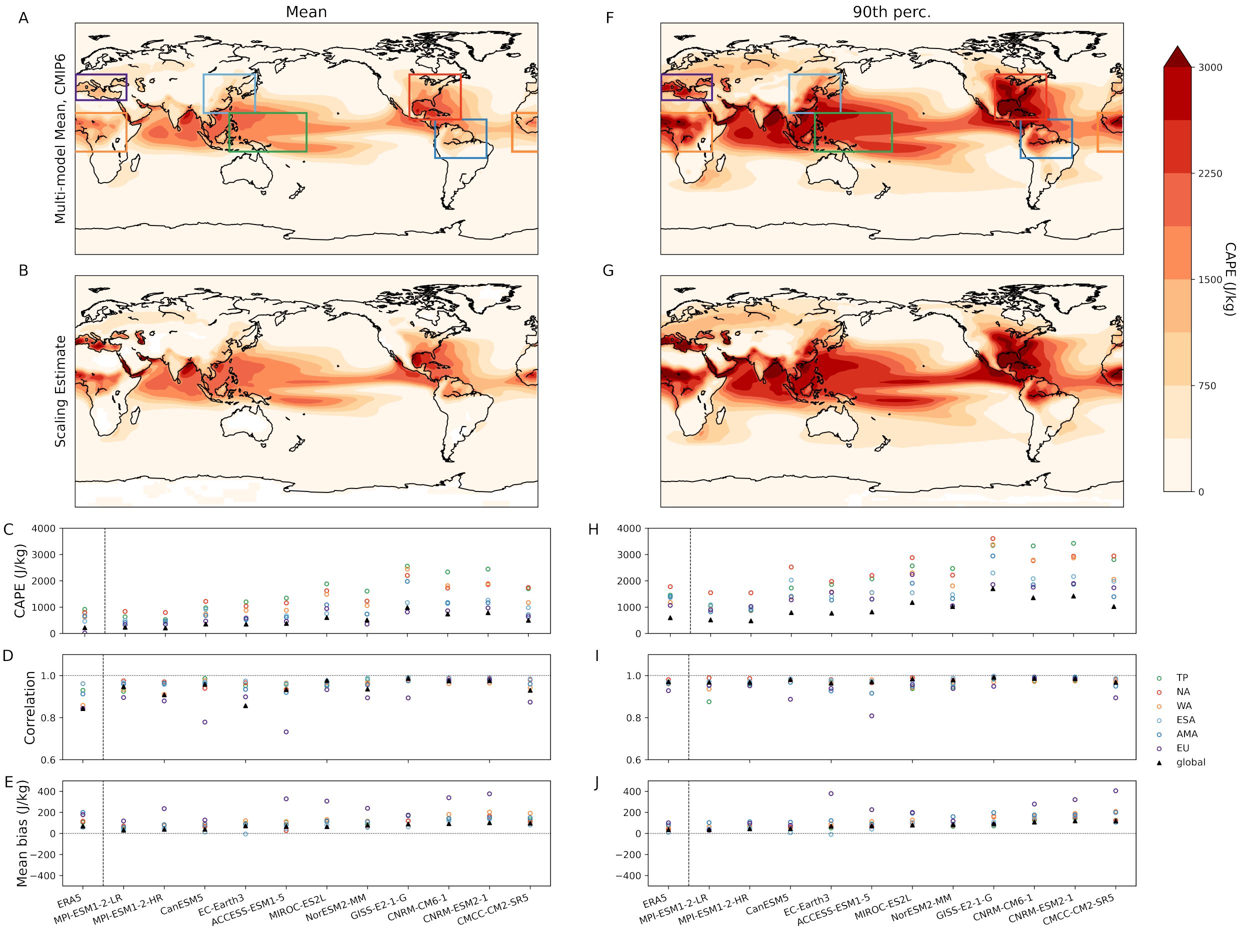

Figure 4.2: Spatial evaluation of (A-E) mean and (F-J) 90th percentile CAPE, 2005–2014 JJA. Ensemble averaged spatial patterns of (top) modeled CAPE and (2nd row) regional CAPE scaling are shown here. The regional CAPE scaling is obtained by deriving convoluted k in 21 x 21 deg boxes (See Table S3.4 for details). The colored boxes correspond to different sub-regions. The contours mark 40% mean bias in CAPE scaling for regions with at least 500 J/kg CAPE. (Last 3 rows) Mean (or 90th percentile) CAPE, spatial correlation, and mean bias (scaling - modeled CAPE) for different regions, with global mean as a black triangle and different sub-regions in different colored circles. Sub-regions are color-coded consistently in all panels, and models are ordered by ascending mean bias in 90th percentile CAPE. Both the magnitude and gradient of CAPE spatial patterns are well-captured by the regional scaling.

The regional scaling captures spatial variations over locations with relatively vigorous convective activities. Here we evaluate fractional biases in reanalysis and CMIP6 models, with all cases where CAPE is larger than 100 J/kg and between 60 N/S. For ERA5, the scaling captures global mean CAPE with a 14.6% overestimation; across CMIP6 models, the fractional bias ranges from 3 to 17% (both ensemble mean and median: 7%). Across different subregions, the scaling works reasonably well and ranges from 7% to 15% for all subregions other than EU (which has mean bias at 306 J/kg, fractional bias by 53%). The scaling biases are substantially smaller compared to model biases (relative to ERA5): taking Eastern U.S. as an example, scaling biases are 5% in ERA5 and 7% in CMIP6; In contrast, the model bias averaged across ensemble members is 76% (ERA5: 794 J/kg, CMIP6 mean: 1401 J/kg). In terms of correlation, the scaling explains ∼95% of spatial variation globally; it also explains more than 90% of regional variation for 55 out of 66 (83%) across all subregions and models. The CNRM models has better correlation across subregions but results in substantial biases for regions like Europe (Mediterranean); the MPI models has better overall performance, striking a lower mean bias and good correlation at the same time. The similar performance for both MPI models suggests that correctly capturing model physics might be more important a factor than model resolution for accurately representing CAPE in models.

Furthermore, we highlight that in an extremely simple global scaling (a single k = 0.52), its spatial pattern matches that of the modeled mean CAPE extremely well, explaining more than 91% of total variances across the globe in ERA5 and 95% when averaged across 11 CMIP6 models; the only drawback is that it fails to reproduce the high CAPE events over land (exceeding 40% fractional bias over regions of high CAPE >500 J/kg, contours in Figure S25), which are critical for impact relevant studies. For sub-regions, the scaling performance remains consistent and robust across models and sub-regions, with high spatial correlations (R^2 above 89%, excluding Europe).

The regional scaling performs similarly well for the 90th percentile CAPE (Figure 4.2, right) – the patterns are retained even for regions with extreme levels of CAPE above 2000 J/kg (e.g., continental U.S., eastern China, and Northern Indian Ocean). The contours of scaling largely follow that of modeled CAPE in all other regions. The scaling explains ∼92% of spatial variation globally (affected by an outlier, otherwise 96%); it also explains more than 90% of regional variation for 60 out of 66 (91%) across all subregions and models. The scaling preserves fundamental physics and preserves a highly consistent free parameter k across mean and 90th percentile CAPE cases (0.52 and 0.56 for CMIP6 ensemble mean). The consistency holds across models, ranging from 0.45 to 0.60 for mean CAPE and 0.49 to 0.64 for 90th percentile CAPE. The consistency suggests we could assume a fixed k for all moderately and strongly convective profiles (e.g., $>$ 100 J/kg CAPE), allowing a consistent CAPE scaling across middle to high percentiles.

4.3.3 Applicability across time scales

A valid scaling should also apply across time scales. Therefore, we examine whether the scaling captures these variations over the Eastern U.S. (Figure 4.3) in the MPI-ESM1-2-LR model [selected following 22]. We assume the “shape of profile” k remains constant at a given time scale; This assumption faces some challenges for the seasonal cycle, where k is larger during the winter and smaller during summer. This is consistent with our understanding that k is a function of dryness – k increases with decreased relative humidity. Fortunately, wintertime means CAPE is much smaller, and the scaling with a constant k across seasons works reasonably well.

The scaling works well for CAPE variations over the mid-latitudes (e.g., Eastern U.S.) on seasonal and inter-annual time scales, with slight overestimation of the seasonal cycles as 787 J/kg versus 620 J/kg modeled and inter-annual variability as 99 versus 82 J/kg modeled (both below $\sim$ 25% fractional bias). However, it severely overestimates the diurnal cycle as 269 J/kg versus 152 J/kg modeled (or a 37% increase from night to day instead of 19%). This overestimation is driven by a compound effect of exaggerated diurnal variations in MSE surplus by 24% and Δz by 10% in the MPI-ESM1-2-LR model. For Western Pacific (see Figure S26), the scaling misses the highs and lows and exhibits too little variation (underestimation by 30%). A raw time series without any temporal averaging is shown in

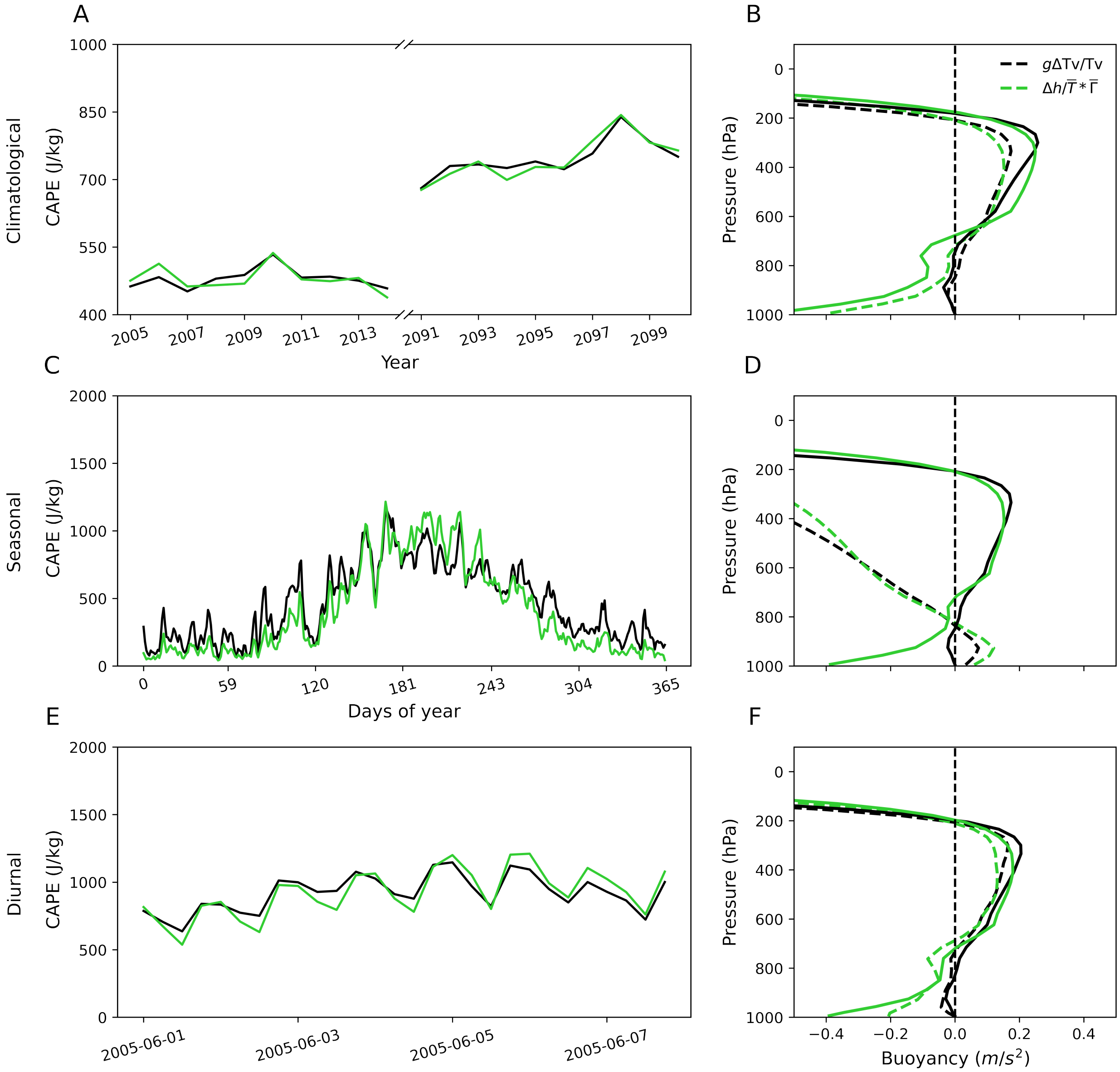

Figure 4.3: Time series and profiles across different time scales for Eastern U.S. in MPI-ESM1-2-LR model: (A-B) climatology; (C-D) seasonal; (E-F) diurnal scales. The left column shows the time series for CAPE (black) and CAPE scaling (limegreen) for each time scale. The monthly mean is shown for climatological, and 6-hourly data are used for seasonal and diurnal. The scaling captures the climatological states and synoptic scale weather episodes; it overestimates the magnitude of diurnal variation. The right column shows averaged profiles above 90th percentile CAPE for each subset. The lines are buoyancy based on (black) virtual temperature and (limegreen) saturated MSE, corresponding to the raw definition of CAPE and our scaling, respectively. The dashed and solid lines are (B) current and future climate, (D) winter and summer, and (F) nighttime and daytime, respectively. The scaling matches the mean profiles and their changes across different time scales.

Figure S27: the scaling performs almost perfectly for diurnal and episodal variations, with R^2 above 0.98 and fractional bias below 10% for all sub-regions evaluated over June 2005.

The scaling can also capture the high CAPE profiles (top 10th percentile) across time scales. The integrated buoyancy profiles based on virtual temperature represent that of raw CAPE, and saturated MSE represents that of our scaling. The mean CAPE from them matches each other extremely well for climatological and diurnal time scales. For summertime in climate states, we have 1145 J/kg for modeled CAPE and 1115 J/kg for scaling; for the end-of-century ssp585 simulation, these values are 1723 and 1911 J/kg. We find similarly good performance when applying the scaling across diurnal time scales: since the high CAPE profiles come pre-dominantly from daytime, we obtain the same numbers for day (1367 and 1371 J/kg); on contrast, we have 1045 and 943 J/kg for summertime night (non-zero because we have substantial coverage of ocean surface; would have much strong diurnal cycle from 493 to 1298 J/kg if restrict to land only). For example, the fractional biases for summertime profiles in both climates are below 11%; the performance drops if we evaluate the mean profile, at ∼20% fractional bias for the current climate.

The performance of scaling drops for wintertime when the atmosphere stability is high (little MSE surplus or CAPE). The modeled CAPE is 65 J/kg, compared to the scaling of 137 J/kg. This difference in performance is mainly an error introduced by the assumption that the "water vapor effect on density" is neglected. In reality, the environmental RH variation is substantial from winter to summer.

4.3.4 Applicability across climate states

A useful and reliable scaling should work for current and future climate states and thus should inform how CAPE would respond in a warmer climate. We quantify mean fractional changes between climate states with the density heatmap and orthogonal distance regression (see Methods and Figure 4.4). The variables are acquired from the model output of current

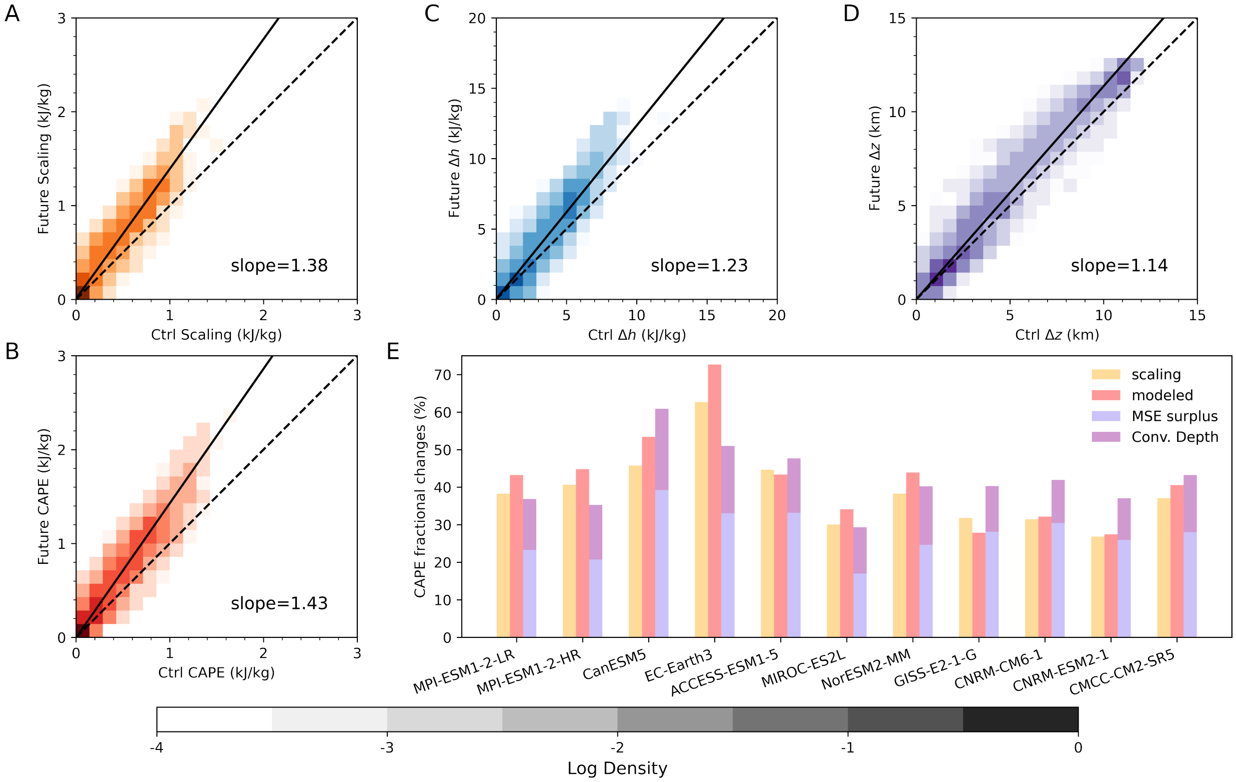

Figure 4.4: Density heatmap of (A) CAPE scaling, (B) CAPE (C) MSE ratio and (D) convective layer depth in current (x-axis) and future (y-axis) space, for MPI-ESM1-2-LR model. The heatmap color bar uses a log10 scale to visualize the occurrence in current–future space. The dashed lines are one-to-one lines, and the solid lines are orthogonal regression lines. Panel (E) shows CAPE fractional changes from models, scaling, and decomposition (contributed by MSE ratio and convective layer depth, respectively). The slopes of the regressed lines represent the fractional changes between climate states, with 1.43 for the CAPE scaling roughly matching 1.38 from the modeled CAPE.

and future climates; these two runs are paired in space and time to create the density heatmap. Mean fractional changes in each variable and the full CAPE scaling are acquired from the slope of the regressions. We find that Eq. 4.1 captures the modeled mean fractional changes well: the scaling increases by 38% in the MPI-ESM1-2-LR model output (39% across all models), compared to the 43% increase in modeled CAPE (42% across all models). The scaling works well in capturing the modeled CAPE changes, with a mean bias by -7.8%. Note that the models diverge in their performance, with two models (ACCESS-ESM1-5 and GISS-E2-1-G) overestimating the fractional change (7.3%) and others underestimating it (-10.6%).

A distinct value of the scaling is its ability to decompose CAPE changes into dominant factors (Figure 4.4E). In this section, we compared the representation of CAPE mean fractional changes in our scaling with CAPE and then evaluated the applicability of the decomposition framework (see Methods). The decomposition results in a 37% increase, derived from a 23% increase in MSE surplus and a 14% increase in Δz – the change predicted by the two terms combined broadly matches that expected from both the modeled CAPE and scaling. The decomposition inexplicitly assumes the terms are uncorrelated – this assumption is acceptable but not perfect. The stacked bar roughly matches that of the modeled in many models (e.g., MPI-ESM1-2-LR 36% instead of 43% modeled, or 38% in scaling), with others having noticeable bias compared to the full scaling (e.g., EC-Earth3 51% instead of 73%). The decomposition works extremely well when averaged across models, yielding the same 42% mean fractional change in CAPE; however, this is a result of canceling errors between the diverging groups of models over- and under-estimating the fractional change, with some members severely underestimating the change (e.g., EC-Earth3 by -16%) and others overestimating it (e.g., GISS-E2-1-G by +15%). Regardless, the scaling works well beyond the current climate, reproducing the expected mean climatological shifts across different CMIP6 models and providing a useful tool to decompose drivers of CAPE changes.

4.3.5 Decomposition of dominant factors across time scales

The decomposition enables physical interpretation for drivers of CAPE changes across space and different temporal scales (Figure S28 for mean and Figure S29 for the 90th percentile). The primary drivers of CAPE can be reduced down to MSE surplus (Δh) and convective layer depth (Δz): the two terms combined explain more than two-thirds of the total variations across space across diurnal, seasonal, or interannual temporal scales. This is true for both the regional evaluations (e.g. Eastern U.S., Table S6) and global evaluation (Table S7). The covariance term between Δh and Δz contributes a substantial portion to the total variance, but a decomposition could still help determine the relative importance of MSE surplus and convective layer depth on CAPE fractional changes on given time scales.

On the climatological time scale, an increase in MSE surplus (surface MSE) following C–C scaling drives the CAPE changes, contributing to 69% of total fractional change. This agrees with the importance of moistening with warming on CAPE changes [35]. On the diurnal time scale, the variations of CAPE are driven by strong temperature swings over the land [29], and MSE surplus contributes to 65% of mean fractional change. On the seasonal time scale, however, the fractional change is driven by convective layer depth changes with insolation patterns, and MSE surplus only contributes to 23% of the total change. The scaling framework provides an intuitive way to explain the underlying physical process of CAPE variations across different time scales.

4.4 Discussion

In this work, we proposed a physical scaling framework of CAPE, explaining its variations across space and diurnal, seasonal, and climatological time scales. The framework has three distinct values compared to previous theoretical models: 1) Only a few physical parameters are needed to construct the scaling, and most of the variation is explained by MSE surplus and convective layer depth; both can be easily derived from vertical profiles without integration. 2) The mathematically constrained "shape of profile" parameter remains highly consistent across CMIP6 models, and the spread holds between mean (0.45–0.60) and 90th percentile CAPE conditions (0.49–0.64) under the current climate; the averaged k also remains constant between climate states (0.52 for current, 0.53 for future). This allows a universally consistent representation of CAPE assuming a constant $k \sim 0.5$. 3) The decomposition makes the physical interpretation of CAPE variations concise and clear, unveiling the physical processes driving CAPE changes across spatial and temporal scales.

This model is not intended to be a full theory of CAPE, in which case it should have a predictive power of CAPE with surface conditions alone and does not require any information on vertical stratification (atmospheric profiles). Our scaling has the potential to become a full theory, with the last piece of the puzzle being how to model MSE surplus with surface conditions alone. That is an intrinsically hard problem since short-term and spatial variations over mid-latitude land are strongly affected by diabatic processes. We think a feasible approach would use a simple energy balance model with lateral boundary conditions prescribed to represent different latitudes and different climate states. The goal of the approach is to understand how differences in MSE surplus and convective layer depth (and other physical parameters) could be obtained given different local and lateral conditions. Nevertheless, our framework provides a practical and straightforward way to understand convective storms in current and future climates. The framework's simplicity enables broader usage across climate and social sciences, potentially fostering stronger connections and collaborations between fields.

4.5 Data Availability Statement

ERA5 reanalysis is available from Copernicus Climate Change Service (C3S, `https://climate.copernicus.eu/climate-reanalysis`). CMIP6 model output is available from the Earth System Grid Federation (ESGF, `https://esgf-node.llnl.gov/projects/cmip6/`).

CHAPTER 5

CONCLUSION

In this thesis, I evaluated CAPE representation in reanalysis and a high-resolution model against radiosonde observations under the current climate. I then examine the minimum information required to project the distributional shift of CAPE under future climate. Finally, I connect the pieces together with a full scaling of CAPE. The scaling works across space and different temporal scales within each and between climate states, providing a straightforward pathway for a better physical understanding of CAPE variations.

The model representation of the convective population and the large-scale parameter CAPE is the foundation for impact attribution under the current climate. We identified low biases in the "high tail" of CAPE distribution across reanalyses and a convection-permitting WRF model with around 200,000 proximity soundings. The reanalyses and WRF CAPE distributions are too narrow, underestimating the occurrence of extreme CAPE values that are most relevant to severe weather events. The modeled SBCAPE is underestimated by 6–10% in the "high tail" above the 95th percentile, while MLCAPE and MUCAPE experience even more severe underestimation by 15–20%. We attribute the sources of these low biases to the underprediction of hot and humid conditions across reanalyses and the WRF model over the Contiguous United States (CONUS). These results emphasize the importance of a realistic planetary boundary layer (PBL) treatment in the models [24, 26]. Beyond the imperfect representation, these errors are likely a result of averaging physical variables, which is pertinent to all model products. The WRF model improves the CAPE representation from its nudging dataset (ERA-Interim) towards ERA5 (which has a higher native resolution), providing the state-of-the-art regional climate simulations we continue our further analysis with.

Convective extremes are expected to increase in intensity and occurrence in a warmer future climate. Regional climate model simulations at convection-permitting scale [e.g., WRF simulations evaluated in the previous chapter, 105] are key to our advanced understanding of these extreme events with climate change. We conclude that CAPE experiences a distributional shift, which is insufficient to be described as a mean fractional change. We find that CAPE is a strong function of MSE surplus; this dependence does not change across climate states, and it also holds across a nudged regional WRF simulation and 11 free-running CMIP6 models. This strong dependence suggests that the distributional shift of CAPE can be readily captured by 3 mean parameter changes: surface T, RH, and mid-tropospheric T. We built the 3-parameter synthetic profiles and confirmed its effectiveness in projecting the full distributional shift. The framework enables a more accurate representation of CAPE changes in the middle to high percentiles and, thus, allows a better understanding of the societal impacts of the convective extremes under climate change.

The CAPE framework based on MSE surplus provides a robust and clear understanding of its changes across climate states. However, our simple scaling shows less consistent dependences on seasonal time scales (winter versus summer) and across latitude bands (within CONUS), presenting limitations in its applicability on small spatial and short time scales – a robust scaling is needed to explain CAPE variations more comprehensively. We derive the full CAPE scaling from two forms, the buoyancy form and the entropy form. The scaling takes only a few physical parameters, and most of the variation can be explained by MSE surplus and convective layer depth. We validate the scaling and confirm its applicability to the full probability density function across space and different temporal scales. In particular, the scaling captures the climatological shifts almost perfectly, predicting a 39% mean CAPE change versus 42% as CMIP6 ensemble mean change. The scaling allows CAPE changes to be decomposed into different driving physical parameters on different time scales, thus providing a straightforward way to understand physical processes driving CAPE variations across space and time.

Although CAPE is a very useful and powerful large-scale parameter for convective extremes, improper interpretation and inference based on CAPE could result in misleading scientific insights into convective weather events. Several limitations should be noted. Firstly, not all potential energy is turned into kinetic energy, and dissipation of moisture is a substantial entropy sink that can't be neglected [96, 95, 119]. Also, CAPE assumes adiabatic ascent; but the entrainment and mixing of dry air from the environment plays a substantial role and needs to be properly addressed [151, 31, 123]. [97] provides an expression to approximate the maximum updraft velocity with entraining CAPE. Still, updraft velocity may be a less relevant parameter than precipitation for societal impact quantification of these extreme events [148]. Future work should map our understanding of large-scale thermodynamic parameters (e.g. CAPE) and their changes to impact-relevant parameters (e.g. precipitation).

Extreme precipitation scaling requires a more accurate representation of thermodynamics and dynamics components than a single condensed metric like CAPE [92, 98]. Furthermore, the impact of these convective extremes is a composite effect of multiple factors, for example: intensity, duration, and size [77, 57, 18]. These storm characteristics can be obtained using storm tracking algorithms and observations or models [19, 20, 135]. Future work should emphasize how large-scale parameters shed light on changes in storm characteristics on synoptic and climatological time scales. The full scaling proposed in this thesis takes us one step closer to a more comprehensive understanding of the societal impacts of convective storms.

REFERENCES

[1] David K. Adams and Enio P. Souza. CAPE and convective events in the southwest during the North American monsoon. *Monthly Weather Review*, 137(1):83–98, 2009.

[2] Vince Agard and Kerry Emanuel. Clausius–Clapeyron scaling of peak CAPE in continental convective storm environments. *Journal of the Atmospheric Sciences*, 74(9):3043–3054, 2017.

[3] John Allen, David Karoly, and Graham Mills. A severe thunderstorm climatology for Australia and associated thunderstorm environments. *Australian Meteorological and Oceanographic Journal*, 61(3):143–158, 2011.

[4] John T. Allen and David J. Karoly. A climatology of Australian severe thunderstorm environments 1979–2011: inter-annual variability and ENSO influence. *International Journal of Climatology*, 34(1):81–97, 2014.

[5] Yuya Baba. Spectral cumulus parameterization based on cloud-resolving model. *Climate Dynamics*, 52(1):309–334, 2019.

[6] P. Bechtold, E. Bazile, F. Guichard, P. Mascart, and E. Richard. A mass-flux convection scheme for regional and global models. *Quarterly Journal of the Royal Meteorological Society*, 127(573):869–886, 2001.

[7] Callyn Bloch, Robert O. Knuteson, Antonia Gambacorta, Nicholas R. Nalli, Jessica Gartzke, and Lihang Zhou. Near-real-time surface-based CAPE from merged hyperspectral IR satellite sounder and surface meteorological station data. *Journal of Applied Meteorology and Climatology*, 58(8):1613–1632, 2019.

[8] William G. Blumberg, Kelton T. Halbert, Timothy A. Supinie, Patrick T. Marsh, Richard L. Thompson, and John A. Hart. SHARPpy: An open-source sounding analysis toolkit for the atmospheric sciences. *Bulletin of the American Meteorological Society*, 98(8):1625–1636, 2017.

[9] Laurens M. Bouwer. Observed and projected impacts from extreme weather events: Implications for loss and damage. In *Loss and Damage from Climate Change: Concepts, Methods and Policy Options*, pages 63–82. Springer International Publishing, 2019.

[10] H. E. Brooks. Severe thunderstorms and climate change. *Atmospheric Research*, 123:129–138, 2013.

[11] Harold E. Brooks, Aaron R. Anderson, Kathrin Riemann, Irina Ebbers, and Heather Flachs. Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. *Atmospheric Research*, 83(2):294–305, 2007.

[12] Harold E Brooks and Jeffrey P Craven. 16.2 A database of proximity soundings for significant severe thunderstorms. In *Extended Abstracts, 21st Conference on Severe Local Storms*, San Antonio, TX, 2002. AMS.

[13] Harold E. Brooks, Charles A. Doswell, and Jeremy Cooper. On the environments of tornadic and nontornadic mesocyclones. *Weather and Forecasting*, 9(4):606–618, 1994.

[14] Harold E Brooks, James W Lee, and Jeffrey P Craven. The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmospheric Research*, 67-68:73–94, 2003.

[15] Andrew Brown, Andrew Dowdy, Todd P. Lane, and Stacey Hitchcock. Types of Severe Convective Wind Events in Eastern Australia. *Monthly Weather Review*, 151(2):419 – 448, 2023.

[16] Matthew J Bunkers, Brian A Klimowski, and Jon W Zeitler. The importance of parcel choice and the measure of vertical wind shear in evaluating the convective environment. In *Extended Abstracts, 21st Conference on Severe Local Storms*, P8.2, San Antonio, TX, 2002. AMS.

[17] C3S. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, 2017. Accessed Sep. 2019, `https://cds.climate.copernicus.eu/cdsapp#!/home`.

[18] Steven C. Chan, Elizabeth J. Kendon, Hayley J. Fowler, Abdullah Kahraman, Julia Crook, Nikolina Ban, and Andreas F. Prein. Large-scale dynamics moderate impact-relevant changes to organised convective storms. *Communications Earth & Environment*, 4(1):1–10, 2023.

[19] Won Chang, Michael L. Stein, Jiali Wang, V. Rao Kotamarthi, and Elisabeth J. Moyer. Changes in spatiotemporal precipitation patterns in changing climate conditions. *Journal of Climate*, 29(23):8355–8376, 2016.

[20] Won Chang, Jiali Wang, Julian Marohnic, V. Rao Kotamarthi, and Elisabeth J. Moyer. Diagnosing added value of convection-permitting regional models using precipitation event identification and tracking. *Climate Dynamics*, 55(1):175–192, 2020.

[21] Alexander B. Charn and Hossein Parishani. Predictive Proxies of Present and Future Lightning in a Superparameterized Model. *Journal of Geophysical Research: Atmospheres*, 126(17):e2021JD035461, 2021.

[22] Daniel R. Chavas and Funing Li. Biases in CMIP6 Historical U.S. Severe Convective Storm Environments Driven by Biases in Mean-State Near-Surface Moist Static Energy. *Geophysical Research Letters*, 49(23):e2022GL098527, 2022.

[23] Jiao Chen, Aiguo Dai, Yaocun Zhang, and Kristen L. Rasmussen. Changes in Convective Available Potential Energy and Convective Inhibition under Global Warming. *Journal of Climate*, 33(6):2025–2050, 2020.

[24] Ariel E. Cohen, Steven M. Cavallo, Michael C. Coniglio, Harold E. Brooks, and Israel L. Jirak. Evaluation of Multiple Planetary Boundary Layer Parameterization Schemes in Southeast U.S. Cold Season Severe Thunderstorm Environments. *Weather and Forecasting*, 32(5):1857–1884, 2017.

[25] Michael C. Coniglio. Verification of RUC 0–1-h Forecasts and SPC Mesoscale Analyses Using VORTEX2 Soundings. *Weather and Forecasting*, 27(3):667–683, 2012.

[26] Michael C. Coniglio, James Correia, Patrick T. Marsh, and Fanyou Kong. Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations. *Weather and Forecasting*, 28(3):842–862, 2013.

[27] Virginia Edith Cortés-Hernández, Feifei Zheng, Jason Evans, Martin Lambert, Ashish Sharma, and Seth Westra. Evaluating regional climate models for simulating sub-daily rainfall extremes. *Climate Dynamics*, 47(5):1613–1628, 2016.

[28] Jeffrey P. Craven, Ryan E. Jewell, and Harold E. Brooks. Comparison between Observed Convective Cloud-Base Heights and Lifting Condensation Level for Two Different Lifted Parcels. *Weather and Forecasting*, 17(4):885–890, 2002.

[29] Aiguo Dai. Global precipitation and thunderstorm frequencies. part II: Diurnal variations. *Journal of Climate*, 14(6):1112–1128, 2001.

[30] Isaac Davis, Funing Li, and Daniel Chavas. Future changes in the vertical structure of severe convective storm environments over the U.S. central Great Plains. *arvix*, 2023.

[31] Wim C. de Rooy, Peter Bechtold, Kristina Fröhlich, Cathy Hohenegger, Harm Jonker, Dmitrii Mironov, A. Pier Siebesma, Joao Teixeira, and Jun-Ichi Yano. Entrainment and detrainment in cumulus convection: an overview. *Quarterly Journal of the Royal Meteorological Society*, 139(670):1–19, 2013.

[32] D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011.

[33] C. Deser, F. Lehner, K. B. Rodgers, T. Ault, T. L. Delworth, P. N. DiNezio, A. Fiore, C. Frankignoul, J. C. Fyfe, D. E. Horton, J. E. Kay, R. Knutti, N. S. Lovenduski, J. Marotzke, K. A. McKinnon, S. Minobe, J. Randerson, J. A. Screen, I. R. Simpson, and M. Ting. Insights from earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, 10(4):277–286, 2020.

[34] Clara Deser, Adam Phillips, Vincent Bourdette, and Haiyan Teng. Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, 38(3):527–546, 2012.

[35] Noah S. Diffenbaugh, Martin Scherer, and Robert J. Trapp. Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proceedings of the National Academy of Sciences*, 110(41):16361–16366, 2013.

[36] Wenhao Dong, Yanluan Lin, Jonathon S. Wright, Yuanyu Xie, Xungang Yin, and Jianping Guo. Precipitable water and CAPE dependence of rainfall intensities in China. *Climate Dynamics*, 52(5):3357–3368, 2019.

[37] Leo J. Donner and Vaughan T. Phillips. Boundary layer control on Convective Available Potential Energy: Implications for cumulus parameterization. *Journal of Geophysical Research: Atmospheres*, 108(D22), 2003.

[38] Leo J. Donner, Charles J. Seman, and Richard S. Hemler. Three-dimensional cloud-system modeling of GATE convection. *Journal of the Atmospheric Sciences*, 56(12):1885–1912, 1999.

[39] Charles A. Doswell. Societal impacts of severe thunderstorms and tornadoes: lessons learned and implications for Europe. *Atmospheric Research*, 67-68:135–152, 2003.

[40] Charles A. Doswell and Erik N. Rasmussen. The effect of neglecting the virtual temperature correction on CAPE calculations. *Weather and Forecasting*, 9(4):625–629, 1994.

[41] Imke Durre, Russell S. Vose, and David B. Wuertz. Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*, 19(1):53–68, 2006.

[42] Imke Durre, Russell S. Vose, and David B. Wuertz. Robust automated quality assurance of radiosonde temperatures. *Journal of Applied Meteorology and Climatology*, 47(8):2081–2095, 2008.

[43] ECMWF. Part i: Observations, chapter 2.3. In *IFS Documentation CY41R2*, pages 31–32. ECMWF, 2016. `https://www.ecmwf.int/node/16646`.

[44] K.A. Emanuel. *Atmospheric Convection, Chap. 6*, pages 169–175. Oxford Univ. Press, 1994.

[45] Kerry Emanuel. On the physics of high CAPE. *Journal of the Atmospheric Sciences*, 80(11):2669–2683, 2023.

[46] Kerry Emanuel and Marja Bister. Moist convective velocity and buoyancy scales. *Journal of the Atmospheric Sciences*, 53(22):3276–3285, 1996.

[47] Todd Emmenegger, Yi-Hung Kuo, Shaocheng Xie, Chengzhu Zhang, Cheng Tao, and J. David Neelin. Evaluating tropical precipitation relations in CMIP6 models with ARM data. *Journal of Climate*, 35(19):6343–6360, 2022.

[48] Clark Evans, Steven J. Weiss, Israel L. Jirak, Andrew R. Dean, and David S. Nevius. An evaluation of paired regional/convection-allowing forecast vertical thermodynamic profiles in warm-season, thunderstorm-supporting environments. *Weather and Forecasting*, 33(6):1547–1566, 2018.

[49] Veronika Eyring, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016.

[50] Zhe Feng, Fengfei Song, Koichi Sakaguchi, and L. Ruby Leung. Evaluation of Mesoscale Convective Systems in Climate Simulations: Methodological Development and Results from MPAS-CAM over the United States. *Journal of Climate*, 34(7):2611–2633, 2021.

[51] David L. A. Flack, Matthew Lehnert, Humphrey W. Lean, and Steve Willington. Characteristics of Diagnostics for Identifying Elevated Convection over the British Isles in a Convection-Allowing Model. *Weather and Forecasting*, 38(7):1079 – 1094, 2023.

[52] G. Fosser, S. Khodayar, and P. Berg. Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Climate Dynamics*, 44(1):45–60, 2015.

[53] M. García-Díez, J. Fernández, L. Fita, and C. Yagüe. Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. *Quarterly Journal of the Royal Meteorological Society*, 139(671):501–514, 2013.

[54] Jessica Gartzke, Robert Knuteson, Grace Przybyl, Steven Ackerman, and Henry Revercomb. Comparison of Satellite-, Model-, and Radiosonde-Derived Convective Available Potential Energy in the Southern Great Plains Region. *Journal of Applied Meteorology and Climatology*, 56(5):1499–1513, 2017.

[55] Victor A. Gensini, Thomas L. Mote, and Harold E. Brooks. Severe-thunderstorm reanalysis environments and collocated radiosonde observations. *Journal of Applied Meteorology and Climatology*, 53(3):742–751, 2013.

[56] A. Gettelman, D. J. Seidel, M. C. Wheeler, and R. J. Ross. Multidecadal trends in tropical Convective Available Potential Energy. *Journal of Geophysical Research: Atmospheres*, 107:ACL 17–1–ACL 17–8, 2002.

[57] Avantika Gori, Ning Lin, Dazhi Xi, and Kerry Emanuel. Tropical cyclone climatology change greatly exacerbates US extreme rainfall–surge hazard. *Nature Climate Change*, 12(2):171–178, 2022.

[58] P. H. Groenemeijer and A. van Delden. Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. *Atmospheric Research*, 83(2):473–487, 2007.

[59] S. Grünwald and H. E. Brooks. Relationship between sounding derived parameters and the strength of tornadoes in Europe and the USA from reanalysis data. *Atmospheric Research*, 100(4):479–488, 2011.

[60] F. Guichard, J. C. Petch, J.-L. Redelsperger, P. Bechtold, J.-P. Chaboureau, S. Cheinet, W. Grabowski, H. Grenier, C. G. Jones, M. Köhler, J.-M. Piriou, R. Tailleux, and M. Tomasini. Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quarterly Journal of the Royal Meteorological Society*, 130(604):3139–3172, 2004.

[61] Leopold Haimberger. Homogenization of radiosonde temperature time series using innovation statistics. *Journal of Climate*, 20(7):1377–1403, 2007.

[62] Leopold Haimberger, Christina Tavolato, and Stefan Sperka. Toward elimination of the warm bias in historic radiosonde temperature records–some new results from a comprehensive intercomparison of upper-air data. *Journal of Climate*, 21(18):4587–4606, 2008.

[63] J.A. Hart and W. Korotky. The SHARP workstation v1.50 users guide. Final Tech. Rep., National Weather Service, NOAA, US. Dept. of Commerce, 1991. [Available from NWS Eastern Region Headquarters, 630 Johnson Ave., Bohemia, NY 11716.].

[64] Isaac M. Held and Brian J. Soden. Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19(21):5686–5699, 2006.

[65] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

[66] Ronald L. Holle and Mary Ann Cooper. Lightning occurrence and social vulnerability. In *Atmospheric Hazards*, chapter 1. IntechOpen, 2016.

[67] D. M. Holley, S. R. Dorling, C. J. Steele, and N. Earl. A climatology of Convective Available Potential Energy in Great Britain. *International Journal of Climatology*, 34(14):3811–3824, 2014.

[68] Song-You Hong, Yign Noh, and Jimy Dudhia. A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9):2318–2341, 2006.

[69] Solomon M Hsiang and Amir S Jina. The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones. Working Paper 20352, National Bureau of Economic Research, 2014.

[70] Michael J. Iacono, Jennifer S. Delamere, Eli J. Mlawer, Mark W. Shephard, Shepard A. Clough, and William D. Collins. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13), 2008.

[71] Robert H. Johns and III Doswell, Charles A. Severe local storms forecasting. *Weather and Forecasting*, 7(4):588–612, 1992.

[72] Rudolf Kaltenböck, Gerhard Diendorfer, and Nikolai Dotzek. Evaluation of thunderstorm indices from ECMWF analyses, lightning data and severe storm reports. *Atmospheric Research*, 93(1):381–396, 2009.

[73] N. Kalthoff, B. Adler, Ch. Barthlott, U. Corsmeier, S. Mobbs, S. Crewell, K. Träumner, Ch. Kottmeier, A. Wieser, V. Smith, and P. Di Girolamo. The impact of convergence zones on the initiation of deep convection: A case study from COPS. *Atmospheric Research*, 93(4):680 – 694, 2009.

[74] Austin T. King and Aaron D. Kennedy. North American Supercell Environments in Atmospheric Reanalyses and RUC-2. *Journal of Applied Meteorology and Climatology*, 58(1):71–92, 2019.

[75] Cornelia Klein, Francis Nkrumah, Christopher M. Taylor, and Elijah A. Adefisan. Seasonality and Trends of Drivers of Mesoscale Convective Systems in Southern West Africa. *Journal of Climate*, 34(1):71–87, 2021.

[76] M. Kunz. The skill of convective parameters and indices to predict isolated and severe thunderstorms. *Natural Hazards and Earth System Sciences*, 7(2):327–342, 2007.

[77] Chun-Chao Kuo, Kai Ernn Gan, Yang Yang, and Thian Yew Gan. Future intensity–duration–frequency curves of Edmonton under climate warming and increased Convective Available Potential Energy. *Climatic Change*, 168(3):30, 2021.

[78] W. Ladwig. WRF-python. Boulder, Colorado: UCAR/NCAR, 2017.

[79] Myong-In Lee, Siegfried D. Schubert, Max J. Suarez, Jae-Kyung E. Schemm, Hua-Lu Pan, Jongil Han, and Soo-Hyun Yoo. Role of convection triggers in the simulation of the diurnal cycle of precipitation over the United States Great Plains in a general circulation model. *Journal of Geophysical Research: Atmospheres*, 113, 2008.

[80] Chiara Lepore, Ryan Abernathey, Naomi Henderson, John T. Allen, and Michael K. Tippett. Future Global Convective Environments in CMIP6 Models. *Earth's Future*, 9(12):e2021EF002277, 2021.

[81] Chiara Lepore, Daniele Veneziano, and Annalisa Molini. Temperature and CAPE dependence of rainfall extremes in the eastern United States. *Geophysical Research Letters*, 42(1):74–83, 2015.

[82] Funing Li and Daniel R. Chavas. Midlatitude continental CAPE is predictable from large-scale environmental parameters. *Geophysical Research Letters*, 48(8):e2020GL091799, 2021.

[83] Changhai Liu, Kyoko Ikeda, Roy Rasmussen, Mike Barlage, Andrew J. Newman, Andreas F. Prein, Fei Chen, Liang Chen, Martyn Clark, Aiguo Dai, Jimy Dudhia, Trude Eidhammer, David Gochis, Ethan Gutmann, Sopan Kurkute, Yanping Li, Gregory Thompson, and David Yates. Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, 49(1):71–95, 2017.

[84] Robert A. Maddox and Charles A. Doswell. An examination of jet stream configurations, 500 mb vorticity advection and low-level thermal advection patterns during extended periods of intense convection. *Monthly Weather Review*, 110(3):184–197, 1982.

[85] Osamu Miyawaki, Tiffany A. Shaw, and Malte F. Jansen. Quantifying energy balance regimes in the modern climate, their link to lapse rate regimes, and their response to warming. *Journal of Climate*, 35(3):1045 – 1061, 2022.

[86] M. W. Moncrieff and M. J. Miller. The dynamics and simulation of tropical cumulonimbus and squall lines. *Quarterly Journal of the Royal Meteorological Society*, 102(432):373–394, 1976.

[87] C. J. Morcrette, K. Van Weverberg, H.-Y. Ma, M. Ahlgrimm, E. Bazile, L. K. Berg, A. Cheng, F. Cheruy, J. Cole, R. Forbes, W. I. Gustafson, M. Huang, W.-S. Lee, Y. Liu, L. Mellul, W. J. Merryfield, Y. Qian, R. Roehrig, Y.-C. Wang, S. Xie, K.-M. Xu, C. Zhang, S. Klein, and J. Petch. Introduction to CAUSES: Description of Weather and Climate Models and Their Near-Surface Temperature Errors in 5 day Hindcasts Near the Southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 123(5):2655–2683, 2018.

[88] Caroline J. Muller, Paul A. O'Gorman, and Larissa E. Back. Intensification of precipitation extremes with warming in a cloud-resolving model. *Journal of Climate*, 24(11):2784–2800, 2011.

[89] Richard B. Neale, Jadwiga H. Richter, and Markus Jochum. The Impact of Convection on ENSO: From a Delayed Oscillator to a Series of Events. *Journal of Climate*, 21(22):5904–5924, 2008.

[90] Stephen W. Nesbitt, Robert Cifelli, and Steven A. Rutledge. Storm morphology and rainfall characteristics of TRMM precipitation features. *Monthly Weather Review*, 134(10):2702–2721, 2006.

[91] Guo-Yue Niu, Zong-Liang Yang, Kenneth E. Mitchell, Fei Chen, Michael B. Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, Mukul Tewari, and Youlong Xia. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, 116(D12), 2011.

[92] Paul A. O'Gorman and Tapio Schneider. The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences*, 106(35):14773–14777, 2009.

[93] Dominique Paquin, Ramón de Elía, and Anne Frigon. Change in North American atmospheric conditions associated with deep convection and severe weather using CRCM4 climate projections. *Atmosphere-Ocean*, 52(3):175–190, 2014.

[94] Raju Pathak, Sandeep Sahany, Saroj Kanta Mishra, and S. K. Dash. Precipitation Biases in CMIP5 Models over the South Asian Region. *Scientific Report*, 9(1):9589, 2019.

[95] Olivier Pauluis. Water Vapor and Mechanical Work: A Comparison of Carnot and Steam Cycles. *Journal of the Atmospheric Sciences*, 68(1):91–102, 2011.

[96] Olivier Pauluis and Isaac M. Held. Entropy Budget of an Atmosphere in Radiative–Convective Equilibrium. Part I: Maximum Work and Frictional Dissipation. *Journal of the Atmospheric Sciences*, 59(2):125–139, 2002.

[97] John M. Peters, Daniel R. Chavas, Chun-Yian Su, Hugh Morrison, and Brice E. Coffer. An Analytic Formula for Entraining CAPE in Midlatitude Storm Environments. *Journal of the Atmospheric Sciences*, 80(9):2165 – 2186, 2023.

[98] S. Pfahl, P. A. O'Gorman, and E. M. Fischer. Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, 7(6):423–427, 2017.

[99] Roger Pielke Jr, Matthew G Burgess, and Justin Ritchie. Plausible 2005–2050 emissions scenarios project between 2 °C and 3 °C of warming by 2100. *Environmental Research Letters*, 17(2):024027, 2022.

[100] Andreas F. Prein, Wolfgang Langhans, Giorgia Fosser, Andrew Ferrone, Nikolina Ban, Klaus Goergen, Michael Keller, Merja Tölle, Oliver Gutjahr, Frauke Feser, Erwan Brisson, Stefan Kollet, Juerg Schmidli, Nicole P. M. van Lipzig, and Ruby Leung. A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53(2):323–361, 2015.

[101] Tomáš Púčik, Pieter Groenemeijer, Anja T. Rädler, Lars Tijssen, Grigory Nikulin, Andreas F. Prein, Erik van Meijgaard, Rowan Fealy, Daniela Jacob, and Claas Teichmann. Future Changes in European Severe Convection Environments in a Regional Climate Model Ensemble. *Journal of Climate*, 30(17):6771–6794, 2017.

[102] Tomáš Púčik, Pieter Groenemeijer, David Rýva, and Miroslav Kolář. Proximity Soundings of Severe and Nonsevere Thunderstorms in Central Europe. *Monthly Weather Review*, 143(12):4805–4821, 2015.

[103] Erik N. Rasmussen and David O. Blanchard. A baseline climatology of sounding-derived supercell and Tornado forecast parameters. *Weather and Forecasting*, 13(4):1148–1164, 1998.

[104] K. L. Rasmussen, A. F. Prein, R. M. Rasmussen, K. Ikeda, and C. Liu. Changes in the convective population and thermodynamic environments in convection-permitting regional climate simulations over the United States. *Climate Dynamics*, 2017.

[105] Roy Rasmussen and Changhai Liu. High Resolution WRF Simulations of the Current and Future Climate of North America, 2017. Accessed Oct. 2019, `https://doi.org/10.5065/D6V40SXP`.

[106] Nilton O. Rennó and Andrew P. Ingersoll. Natural Convection as a Heat Engine: A Theory for CAPE. *Journal of Atmospheric Sciences*, 53(4):572 – 585, 1996.

[107] Kathrin Riemann-Campe, Klaus Fraedrich, and Frank Lunkeit. Global climatology of Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in ERA-40 reanalysis. *Atmospheric Research*, 93(1):534–545, 2009.

[108] David M. Romps. Response of tropical precipitation to global warming. *Journal of the Atmospheric Sciences*, 68(1):123–138, 2011.

[109] David M. Romps. Clausius–Clapeyron scaling of CAPE from analytical solutions to RCE. *Journal of the Atmospheric Sciences*, 73(9):3719–3737, 2016.

[110] David M. Romps. Evaluating the future of lightning in cloud-resolving models. *Geophysical Research Letters*, 46(24):14863–14871, 2019.

[111] David M. Romps, Jacob T. Seeley, David Vollaro, and John Molinari. Projected increase in lightning strikes in the United States due to global warming. *Science*, 346(6211):851–854, 2014.

[112] Harald Rybka, Ulrike Burkhardt, Martin Köhler, Ioanna Arka, Luca Bugliaro, Ulrich Görsdorf, Ákos Horváth, Catrin I. Meyer, Jens Reichardt, Axel Seifert, and Johan Strandgren. The behavior of high-CAPE (Convective Available Potential Energy) summer convection in large-domain large-eddy simulations with ICON. *Atmospheric Chemistry and Physics*, 21(6):4285–4318, 2021.

[113] Russ S. Schumacher, Deanna A. Hence, Stephen W. Nesbitt, Robert J. Trapp, Karen A. Kosiba, Joshua Wurman, Paola Salio, Martin Rugna, Adam C. Varble, and Nathan R. Kelly. Convective-Storm Environments in Subtropical South America from High-Frequency Soundings during RELAMPAGO-CACTI. *Monthly Weather Review*, 149(5):1439 – 1458, 2021.

[114] Kevin Schwarzwald, Andrew Poppick, Maria Rugenstein, Jonah Bloch-Johnson, Jiali Wang, David McInerney, and Elisabeth J. Moyer. Changes in future precipitation mean and variability across scales. *Journal of Climate*, 34(7):2741–2758, 2021.

[115] Jacob T. Seeley and David M. Romps. Why does tropical Convective Available Potential Energy (CAPE) increase with warming? *Geophysical Research Letters*, 42(23):10429–10437, 2015.

[116] Martin S. Singh, Zhiming Kuang, Eric D. Maloney, Walter M. Hannah, and Brandon O. Wolding. Increasing potential for intense tropical and subtropical thunderstorms under global warming. *Proceedings of the National Academy of Sciences*, 114(44):11657–11662, 2017.

[117] Martin S. Singh and Paul A. O'Gorman. Influence of entrainment on the thermal stratification in simulations of Radiative-Convective Equilibrium. *Geophysical Research Letters*, 40(16):4398–4403, 2013.

[118] Martin S. Singh and Paul A. O'Gorman. Increases in moist-convective updraught velocities with warming in Radiative-Convective Equilibrium: Increases in updraught velocities with warming. *Quarterly Journal of the Royal Meteorological Society*, 141(692):2828–2838, 2015.

[119] Martin S. Singh and Morgan E O'Neill. The climate system and the second law of thermodynamics. *Reviews of Modern Physics*, 94:015001, 2022.

[120] Adam B. Smith. U.S. Billion-dollar Weather and Climate Disasters, 1980 - present (NCEI Accession 0209268), 2020. Accessed Jun. 2022.

[121] Fengfei Song and Guang J. Zhang. Improving trigger functions for convective parameterization schemes using GOAmazon observations. *Journal of Climate*, 30(21):8711–8726, 2017.

[122] Fengfei Song and Guang J. Zhang. Understanding and improving the scale dependence of trigger functions for convective parameterization using cloud-resolving model data. *Journal of Climate*, 31(18):7385–7399, 2018.

[123] Kenta Sueki and Hiroshi Niino. Toward better assessment of tornado potential in typhoons: Significance of considering entrainment effects for CAPE. *Geophysical Research Letters*, 43(24):12597–12604, 2016.

[124] E. Suhas and Guang J. Zhang. Evaluation of trigger functions for convective parameterization schemes using observations. *Journal of Climate*, 27(20):7647–7666, 2014.

[125] Xuguang Sun, Ming Xue, Jerald Brotzge, Renee A. McPherson, Xiao-Ming Hu, and Xiu-Qun Yang. An evaluation of dynamical downscaling of Central Plains summer precipitation using a WRF-based regional climate model at a convection-permitting 4 km resolution. *Journal of Geophysical Research: Atmospheres*, 121(23):13801–13825, 2016.

[126] Mateusz Taszarek, John T. Allen, Tomáš Púčik, Kimberly A. Hoogewind, and Harold E. Brooks. Severe Convective Storms across Europe and the United States. Part II: ERA5 Environments Associated with Lightning, Large Hail, Severe Wind, and Tornadoes. *Journal of Climate*, 33(23):10263–10286, 2020.

[127] Mateusz Taszarek, Harold E. Brooks, Bartosz Czernecki, Piotr Szuster, and Krzysztof Fortuniak. Climatological Aspects of Convective Parameters over Europe: A Comparison of ERA-Interim and Sounding Data. *Journal of Climate*, 31(11):4281–4308, 2018.

[128] Mateusz Taszarek, Natalia Pilguj, John T. Allen, Victor Gensini, Harold E. Brooks, and Piotr Szuster. Comparison of convective parameters derived from ERA5 and MERRA2 with rawinsonde data over Europe and North America. *Journal of Climate*, -1:1–55, 2020.

[129] Gregory Thompson and Trude Eidhammer. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the Atmospheric Sciences*, 71(10):3636–3658, 2014.

[130] Richard L. Thompson, Roger Edwards, John A. Hart, Kimberly L. Elmore, and Paul Markowski. Close Proximity Soundings within Supercell Environments Obtained from the Rapid Update Cycle. *Weather and Forecasting*, 18(6):1243–1261, 2003.

[131] M. Tiedtke. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117(8):1779–1800, 1989.

[132] Michael K. Tippett, Adam H. Sobel, and Suzana J. Camargo. Association of U.S. tornado occurrence with monthly environmental parameters. *Geophysical Research Letters*, 39(2), 2012.

[133] Robert J. Trapp, Noah S. Diffenbaugh, Harold E. Brooks, Michael E. Baldwin, Eric D. Robinson, and Jeremy S. Pal. Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proceedings of the National Academy of Sciences*, 104(50):19719–19723, 2007.

[134] Robert J. Trapp, Noah S. Diffenbaugh, and Alexander Gluhovsky. Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophysical Research Letters*, 36(1), 2009.

[135] Philip Tuckman, Vince Agard, and Kerry Emanuel. Evolution of convective energy and inhibition before instances of large CAPE. *Monthly Weather Review*, 151(1):321–338, 2023.

[136] Yi-Chi Wang, Hua-Lu Pan, and Huang-Hsiung Hsu. Impacts of the triggering function of cumulus parameterization on warm-season diurnal rainfall cycles at the atmospheric radiation measurement Southern Great Plains site. *Journal of Geophysical Research: Atmospheres*, 120(20):10681–10702, 2015.

[137] Ziwei Wang, James A. Franke, Zhenqi Luo, and Elisabeth J. Moyer. Reanalyses and a high-resolution model fail to capture the "high tail" of CAPE distributions. *Journal of Climate*, 34(21):8699–8715, 2021.

[138] Ziwei Wang and Elisabeth J. Moyer. Robust Relationship between Midlatitudes CAPE and Moist Static Energy Surplus in Present and Future Simulations. *Geophysical Research Letters*, 50(14):e2023GL104163, 2023.

[139] Patrick Willems. Impacts of climate change on rainfall extremes and urban drainage systems. *Water Intelligence Online*, 11, 2012.

[140] Earle Williams and Nilton Renno. An analysis of the conditional instability of the tropical atmosphere. *Monthly Weather Review*, 121(1):21–36, 1993.

[141] M. Wu and J.-E. Lee. Thresholds for Atmospheric Convection in Amazonian Rainforests. *Geophysical Research Letters*, 46(16):10024–10033, 2019.

[142] Shaocheng Xie, Yi-Chi Wang, Wuyin Lin, Hsi-Yen Ma, Qi Tang, Shuaiqi Tang, Xue Zheng, Jean-Christophe Golaz, Guang J. Zhang, and Minghua Zhang. Improved diurnal cycle of precipitation in E3SM with a revised convective triggering function. *Journal of Advances in Modeling Earth Systems*, 11(7):2290–2310, 2019.

[143] Shaocheng Xie and Minghua Zhang. Impact of the convection triggering function on single-column model simulations. *Journal of Geophysical Research: Atmospheres*, 105:14983–14996, 2000.

[144] Lujun Xu, Huizhi Liu, Qun Du, and Xiangde Xu. The assessment of the planetary boundary layer schemes in WRF over the central Tibetan Plateau. *Atmospheric Research*, 230:104644, 2019.

[145] Ben Yang, Yang Zhou, Yaocun Zhang, Anning Huang, Yun Qian, and Lujun Zhang. Simulated precipitation diurnal cycles over East Asia using different CAPE-based convective closure schemes in WRF model. *Climate Dynamics*, 50(5):1639–1658, 2018.

[146] J.-I. Yano, M. Bister, Ž. Fuchs, L. Gerard, V. T. J. Phillips, S. Barkidija, and J.-M. Piriou. Phenomenology of convection-parameterization closure. *Atmospheric Chemistry and Physics*, 13(8):4111–4131, 2013.

[147] Bing Ye, Anthony D. Del Genio, and Kenneth K-W. Lo. CAPE variations in the current climate and in a climate change. *Journal of Climate*, 11(8):1997–2015, 1998.

[148] M. J. B. Zeppel, J. V. Wilks, and J. D. Lewis. Impacts of extreme precipitation and seasonal changes in precipitation on plants. *Biogeosciences*, 11(11):3083–3093, 2014.

[149] G. J. Zhang and Norman A. McFarlane. Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate centre general circulation model. *Atmosphere-Ocean*, 33(3):407–446, 1995.

[150] Guang J. Zhang. Convective quasi-equilibrium in midlatitude continental environment and its effect on convective parameterization. *Journal of Geophysical Research: Atmospheres*, 107(D14):ACL 12–1–ACL 12–16, 2002.

[151] Guang J. Zhang. Effects of entrainment on Convective Available Potential Energy and closure assumptions in convection parameterization. *Journal of Geophysical Research: Atmospheres*, 114(D7), 2009.

[152] Yi Zhang and William R. Boos. An upper bound for extreme temperatures over midlatitude land. *Proceedings of the National Academy of Sciences*, 120(12):e2215278120, 2023.

[153] Dong Zheng, Yijun Zhang, Qing Meng, Luwen Chen, and Jianru Dan. Climatology of lightning activity in South China and its relationships to precipitation and Convective Available Potential Energy. *Advances in Atmospheric Sciences*, 33(3):365–376, 2016.

[154] Wenyu Zhou and Shang-Ping Xie. A conceptual spectral plume model for understanding tropical temperature profile and convective updraft velocities. *Journal of the Atmospheric Sciences*, 76(9):2801–2814, 2019.

APPENDIX

S1 Supplemental Materials for Chapter 2

S1.1 CAPE distributions and biases

Table S1: Mean values of each type of CAPE (unit: J/kg) from the four datasets considered in this study. Summertime summary statistics are shown and are not paired to IGRA. We show averages both with and without zeroes included. The bottom row is corrected SBCAPE, recomputed with surface values replaced with those measured by radiosondes. Note that mean values of the WRF model and reanalysis CAPE are slightly larger than in radiosondes, even though these data products substantially underpredict the "high tail" of CAPE distributions. Overly narrow CAPE distributions may still show reasonable mean values.

		IGRA	ERAI	ERA5	WRF
SBCAPE	w/ zeroes	680	696	724	686
	w/o zeroes	1064	1124	1128	1140
MUCAPE	w/ zeroes	1058	829	856	759
	w/o zeroes	1368	1214	1225	1661
MLCAPE	w/ zeroes	562	446	473	441
	w/o zeroes	912	839	848	834
Corrected SBCAPE	w/ zeroes	680	641	679	653
	w/o zeroes	1064	1088	1089	1107

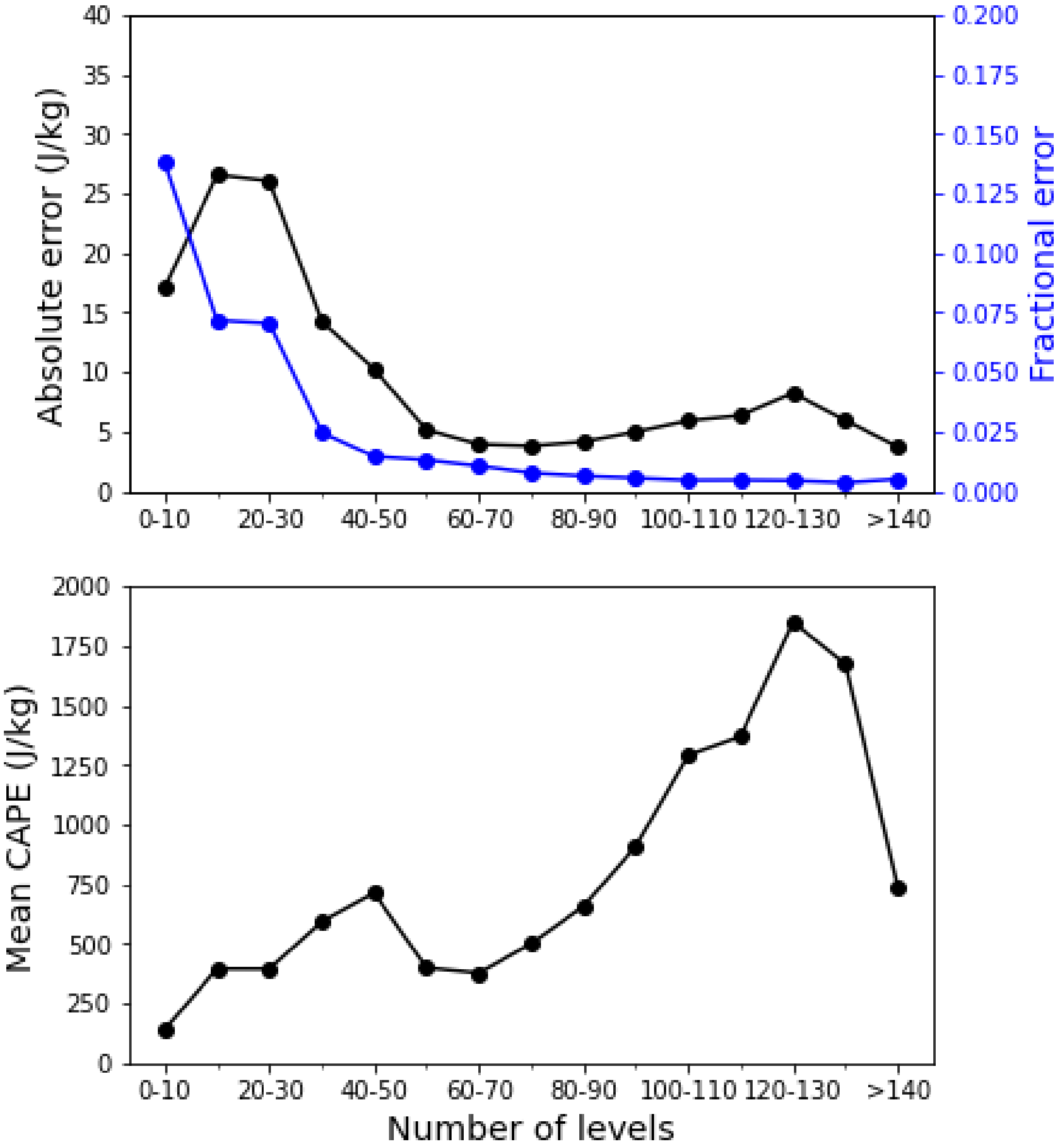

Figure S1: (Top) Sensitivity of interpolation error to the number of levels in raw IGRA profiles. Both absolute error (black, left axis) and fractional error (blue, right axis) are shown. (Bottom) Mean CAPE across number of levels in raw profiles. The lower CAPE from 50–80 levels is responsible to the concave shape of the absolute difference around those levels.

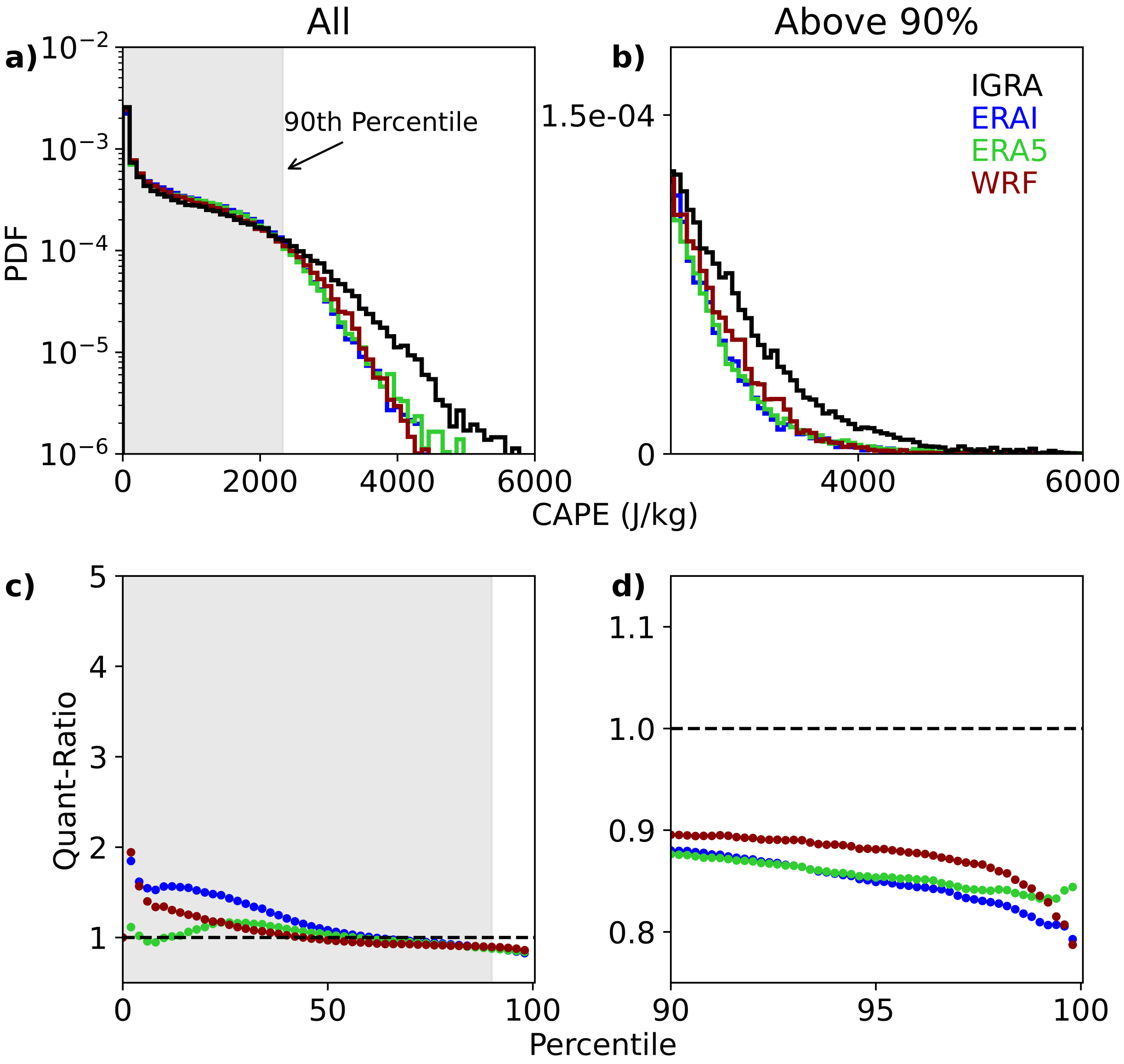

Figure S2: As in Figure 2.2, but for MLCAPE instead of SBCAPE. Points with zero CAPE are excluded from the analysis (39–47% of the datasets). PDFs are cut off at 6000 J/kg on the x-axis, as less than 0.02% of all points lie above the limit. For IGRA, the 90th percentile is 2331 J/kg, the 95th is 2832 J/kg, and the 97.5th is about 3260 J/kg. Biases in the distribution and mean of MLCAPE are midway between those in SBCAPE and MUCAPE.

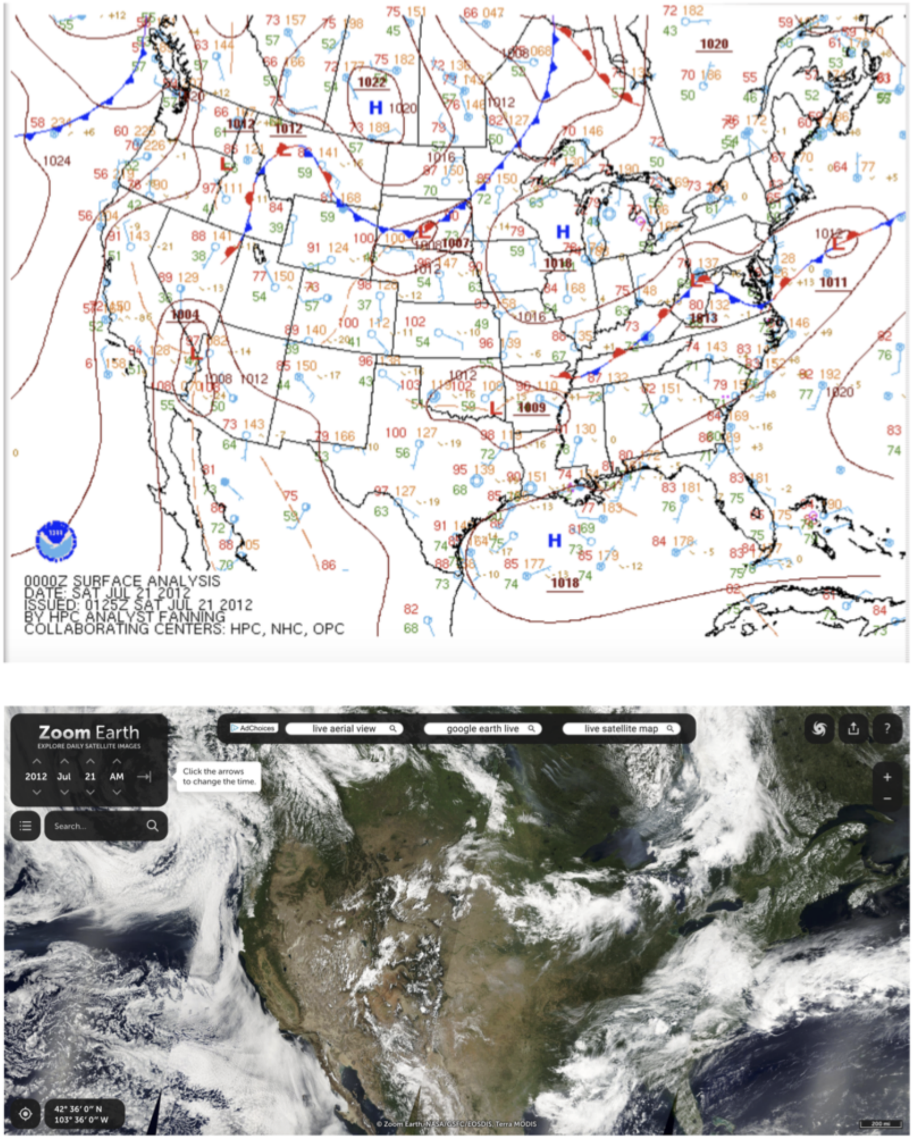

Figure S3: The example day of July 21, 2012, shown as (**top**) a weather map from the National Weather Service (NWS), and (**bottom**) a satellite image from Terra MODIS visualized by Zoom Earth. Compare to manuscript Figure 2.4. Images are available online at $https://www.wpc.ncep.noaa.gov/archives/web\_pages/sfc/sfc\_archive\_maps.php?arcdate = 07/21/2012\&selmap = 2012072100\&maptype = namussfc$, $https://zoom.earth/\#view = 42.6, -103.6, 4.54z/date = 2012-07-21, am$. In this time period a stationary front in the Southeast U.S. brings high moisture and cloud cover to the region. Extreme SBCAPE occurs at the edge of this zone, where moist air meets the high temperatures of the Central U.S.

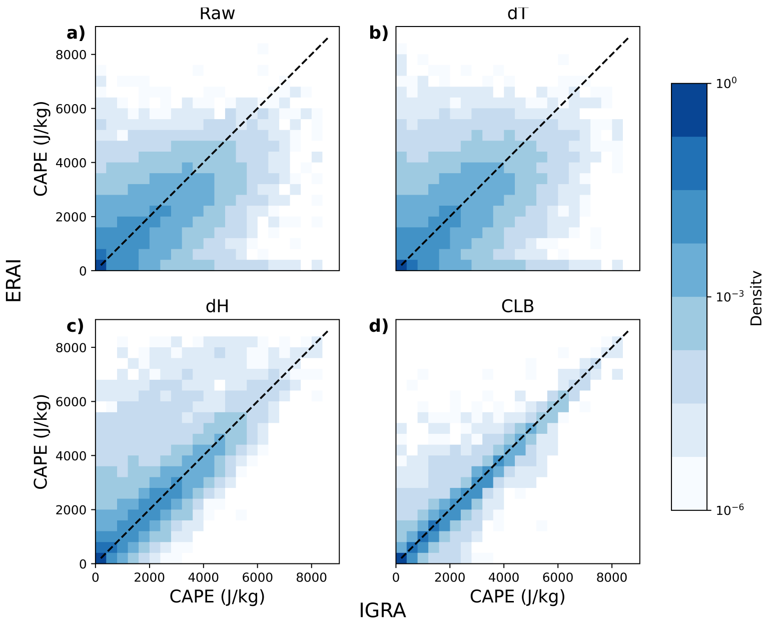

Figure S4: As in manuscript Figure 2.5, but for ERAI instead of WRF.

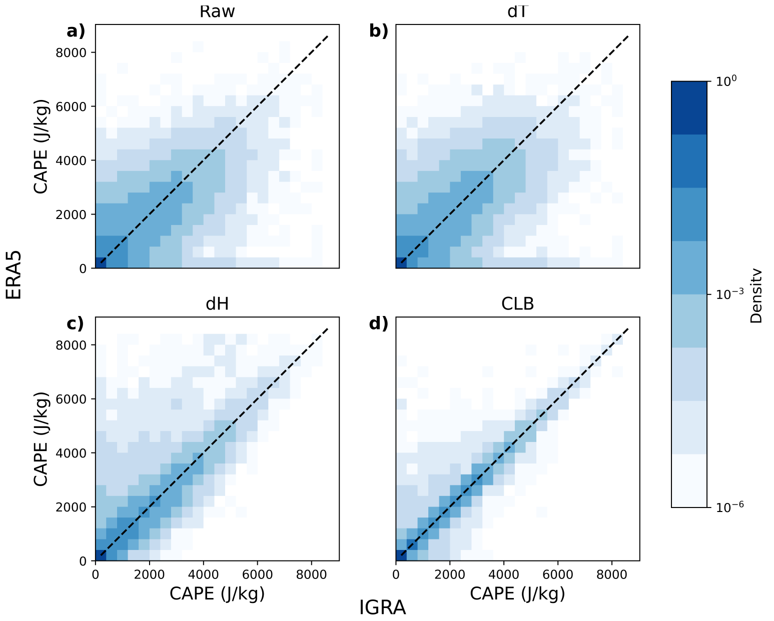

Figure S5: As in manuscript Figure 2.5, but for ERA5 instead of WRF.

Table S2: Improvement of the SBCAPE match to observations by correcting surface values: table shows Pearson-r for the correlation of CAPE from different datasets against radiosondes, using different surface level corrections. Top to bottom rows show values for raw data, surface temperature and pressure corrected, surface humidity and pressure corrected, and with all surface values corrected. All correlation coefficients pass the two-sided test at 0.1% significance level.

	ERAI	ERA5	WRF
Raw	0.80	0.83	0.68
dT	0.81	0.84	0.65
dH	0.93	0.94	0.91
All	0.99	0.99	0.99

S1.2 CAPE as a function of surface T, H

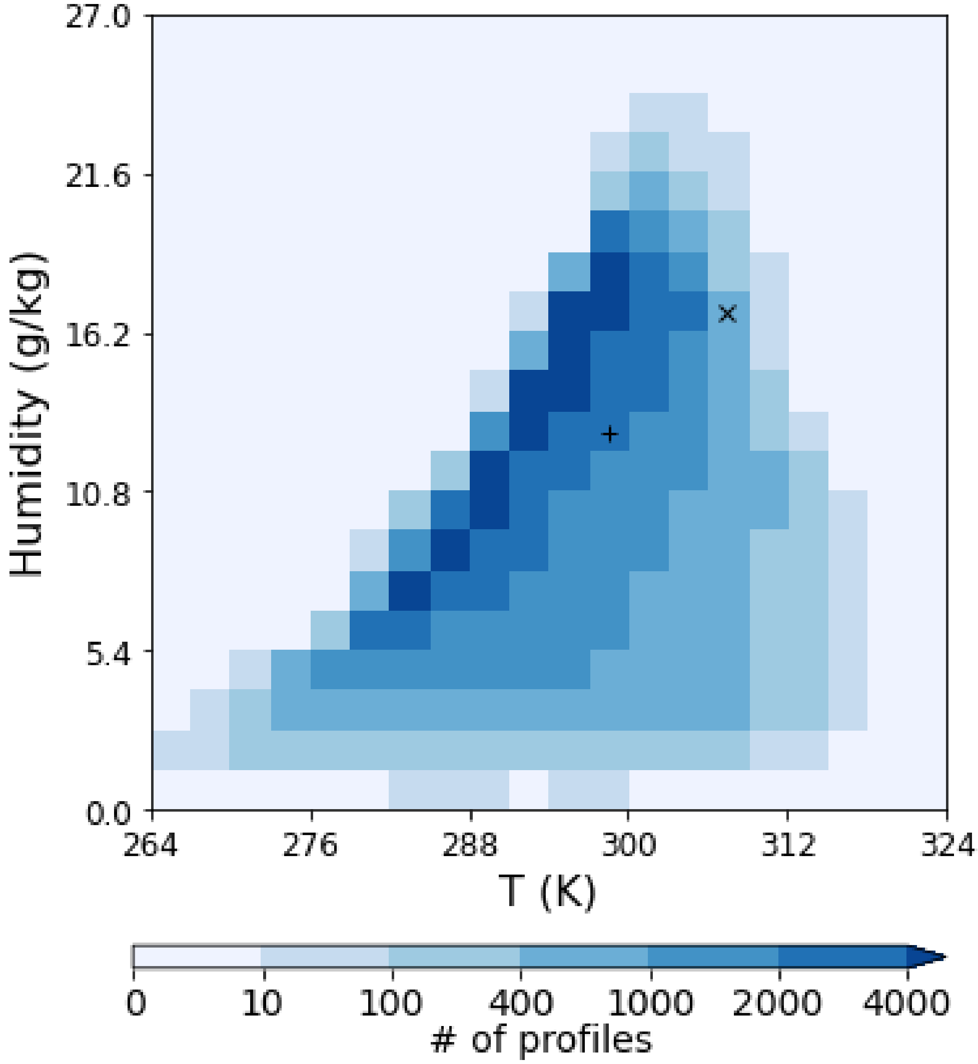

Figure S6: Occurrence of radiosonde records (IGRA) in T–H space. The total number of profiles is 199,787. The colorbar is non-linear in order to highlight the bins with fewer samples in it. Note that the density and absolute occurrence plots are identical other than a scaling factor.

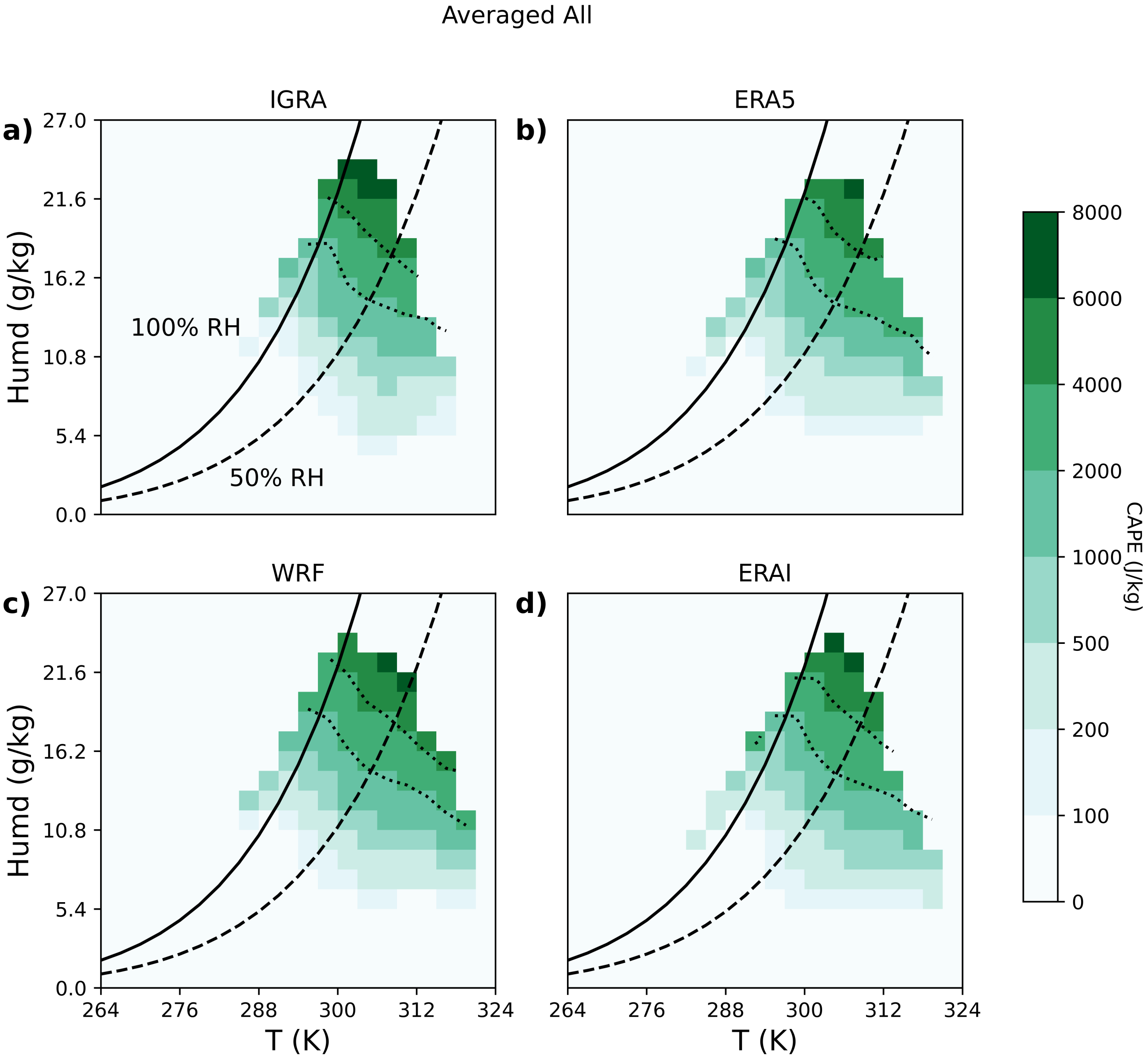

Figure S7: As in manuscript Figure 2.6, but for all datasets. Panels **a–d)** are IGRA, ERA5, WRF, and ERAI, respectively. Contours correspond to 2000 and 4000 J/kg. All datasets show similar dependence of SBCAPE on surface temperature and humidity, but IGRA radiosondes sample more extreme (hot, humid) conditions.

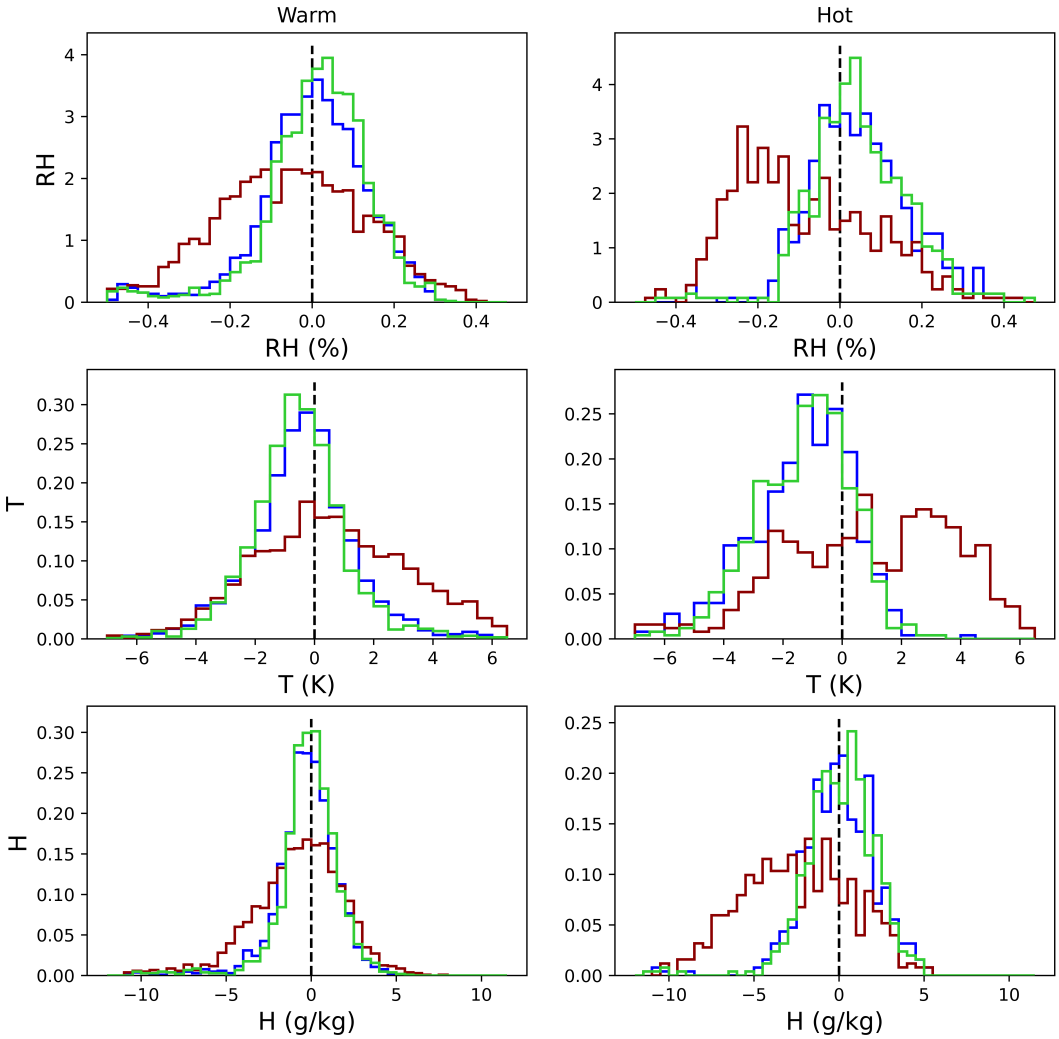

Figure S8: Distributions of errors in reanalyses and the WRF model surface values for the 'warm' and 'hot' bins defined in main text Figure 2.7. Figure shows reanalyses and model errors in profiles assigned to these T–H bins based on their radiosonde values. Mean biases are relatively small, especially in reanalyses. As is the norm, reanalyses are slightly too cold, and WRF is too hot and dry. The distributions of errors are, in most cases, reasonably symmetrical, and mean biases are small relative to standard deviations: for example, ERA5 specific humidity in the 'hot' bin has a mean bias of .007 g/kg but an error standard deviation of 2.3 g/kg. Mean biases, skew in distributions, and standard deviations are largest in WRF. Complete values of mean and error standard deviations are given in Table S3.

Table S3: Mean and standard deviation (in parentheses) of errors in reanalysis and the WRF model in RH, temperature, and specific humidity, all computed against paired radiosonde soundings as in Figure S18. Values are given for (**top, middle**) the 'hot' and 'warm' bins from Figure S8 and manuscript Figure 2.7 and (**bottom**) as averages over all bins in the summertime. Standard deviations of profile errors are relatively constant in all cases, presumably because the weather events that produce strong gradients occur in a wide variety of environments. Biases are stronger in the 'hot' bin, which samples warm, humid conditions associated with extreme CAPE.

		ERAI	ERA5	WRF
'Hot' bin	RH (%)	4.7 (11.8)	4.6 (11.1)	-8.8 (16.0)
	T (K)	-1.4 (1.9)	-1.3 (1.8)	0.9 (3.0)
	H (g/kg)	-0.06 (2.3)	0.005 (2.3)	-2.5 (3.4)
'Warm' bin	RH (%)	0 (12.8)	1.1 (12.3)	-5.7 (17.9)
	T (K)	-0.2 (2.5)	-0.3 (2.4)	1.0 (3.4)
	H (g/kg)	-0.3 (1.9)	-0.1 (1.8)	-0.8 (2.7)
Average	RH (%)	-1.3 (12.6)	-0.8 (11.9)	-6.8 (16.7)
	T (K)	-0.01 (2.7)	0.09 (2.3)	1.2 (2.9)
	H (g/kg)	-0.02 (1.7)	0.06 (1.6)	-0.4 (2.3)

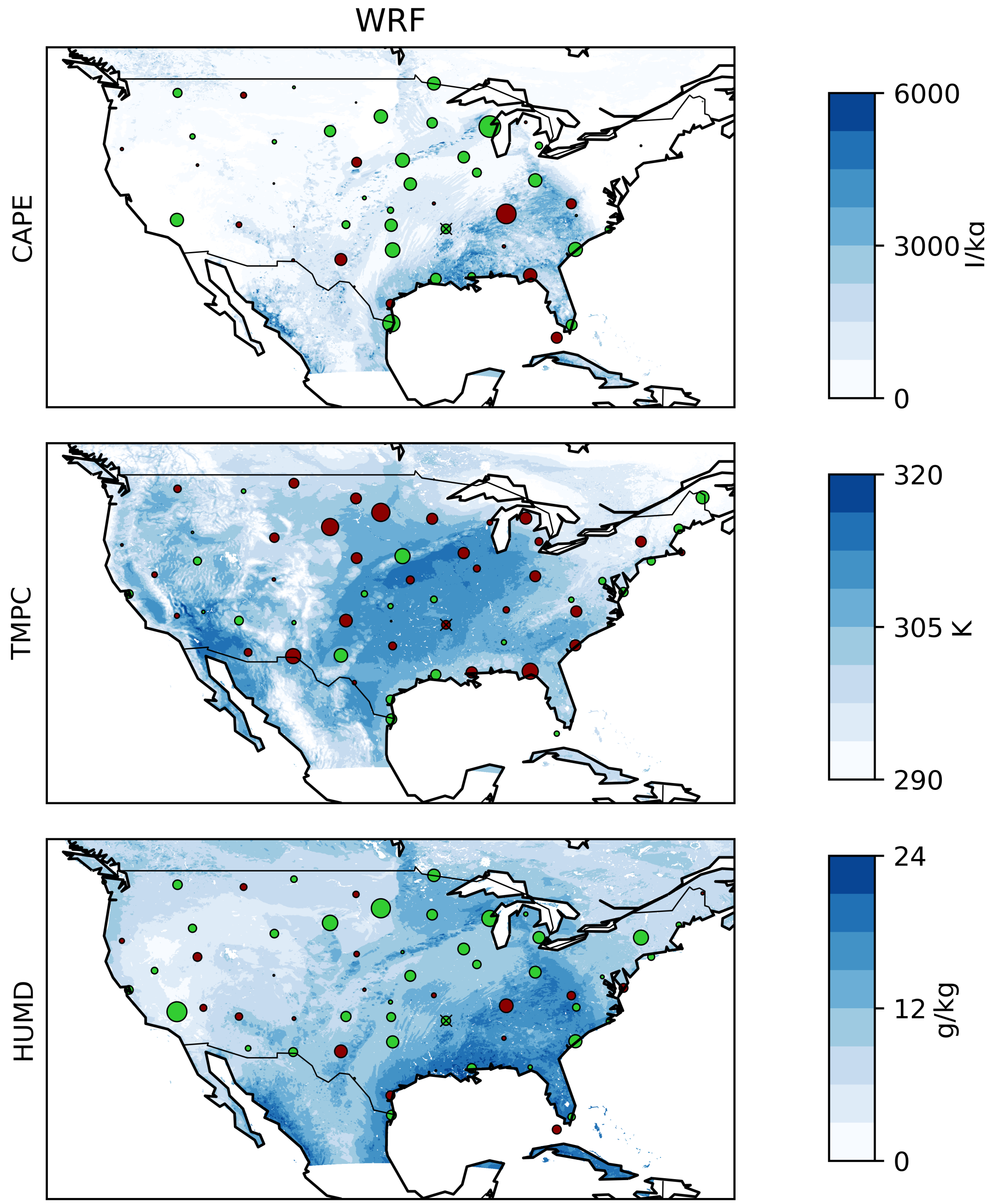

Figure S9: As in manuscript Figure 2.4, but for July 26th, 2012 instead of July 21st, 2012. Again, the hot and dry bias in the WRF model is manifested in Central U.S., which leads to the underestimation in CAPE.

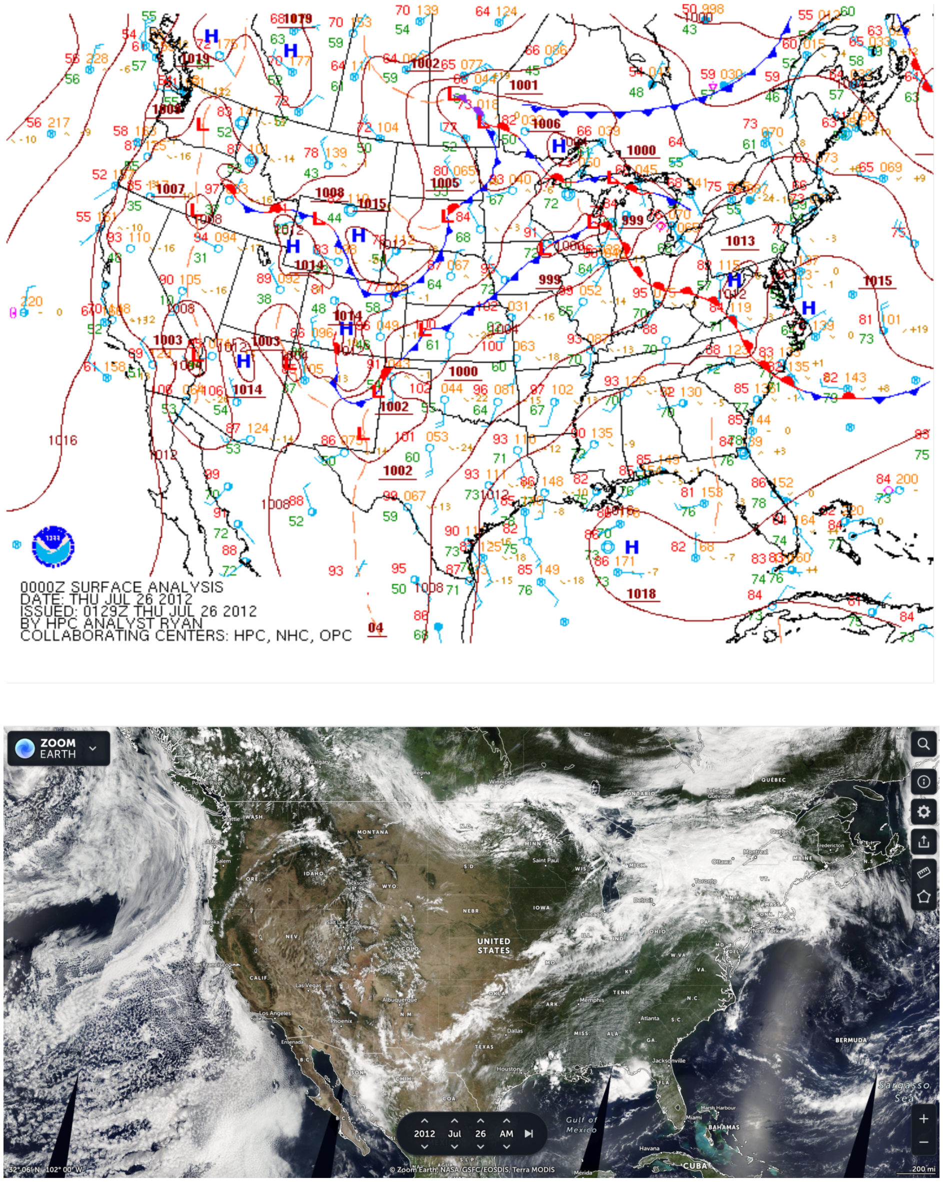

Figure S10: The example day of July 26, 2012 shown as (**top**) a weather map from the National Weather Service (NWS), and (**bottom**) a satellite image from Terra MODIS visualized by Zoom Earth. Compare to Figure S9 Images are available online at $https : //www.wpc.ncep.noaa.gov/archives/web\_pages/sfc/sfc\_archive\_maps.php?arcdate = 07/26/2012\&selmap = 2012072600\&maptype = namussfc, https : //zoom.earth/\#view = 39, -99.9, 4.54z/date = 2012 - 07 - 26, am$. During this time period, a strong cold front in the Midwest U.S. brings cloud system to the region.

S2 Supplemental Materials for Chapter 3

S2.1 Derivation of CAPE-MSE surplus framework

In this section we show the background for the framework in our manuscript. We repeat the derivation from Emanuel, *Atmospheric Convection* (1994) that restates CAPE as a function of pseudo-entropy, show how this can be approximated as a linear dependence in pseudo-enthalpy (moist static energy), and finally demonstrate that the error introduced by the core assumption required in [44, hereafter E94] – that virtual temperature corrections can be ignored – is relatively minor and considerably smaller than that in an alternative CAPE framework.

Emanuel (1994) derivation

We start from the definition of CAPE in pressure coordinates:

$$CAPE = \int_{p_n}^{p_i} (\alpha_p - \alpha_a) dp \tag{5.1}$$

where α_p and α_a are the volume per mass for air in the parcel and the environment, respectively. Because CAPE is a positive quantity, the integration is from low to high pressure, i.e. from the top of a convective event (p_n) to the level of convective initiation (p_i).

The rising parcel will be saturated, and changes in its volume per mass can be divided into two terms, separating the effects of its saturation pseudo-entropy s^* (which is independent of moisture) and of its actual moisture content. That is:

$$\Delta\alpha = \left.\frac{\partial\alpha}{\partial s^*}\right|_r \Delta s^* + \left.\frac{\partial\alpha}{\partial r}\right|_{s^*} \Delta r \tag{5.2}$$

If we treat the environment as also saturated – acceptable if the effect of moisture on density is small – then it can be similarly decomposed and CAPE can be written as:

$$CAPE = \int_{p_n}^{p_i} \left(\frac{\partial \alpha}{\partial s^*}(s_p^* - s_a^*) + \frac{\partial \alpha}{\partial r}(r_p - r_a) \right) \cdot dp \tag{5.3}$$

The volume per mass of dry air (α_d) can be approximated as $\alpha_d = \frac{\alpha}{1+r}$. where r is the mass mixing ratio of water vapor, typically 0.01 or less. Emanuel then makes the further assumption that the buoyancy effects of this water vapor r (the virtual temperature effect) can be neglected entirely so that the second term in Equation 5.3 vanishes and in the first term α is replaced by α_d. This yields Eq. (6.4.2a) in Emanuel (1994):

$$CAPE \approx \int_{p_n}^{p_i} \frac{\partial \alpha_d}{\partial s^*}(s_p^* - s_a^*) \cdot dp \tag{5.4}$$

The neglect of virtual temperature effects for both parcel and environment produces a slight net underestimation of derived CAPE, but the distortion is smaller than in other approximate CAPE frameworks and is compensated for by the empirical regression coefficient. See the discussion at the end of the section and Figure S11. The Maxwell relationship $(\frac{\partial \alpha}{\partial s})_p = (\frac{\partial T}{\partial p})_s$ allows converting the integration coordinate in Equation 5.4 from pressure to temperature:

$$\begin{aligned} CAPE &= \int_{p_n}^{p_i} \frac{\partial T}{\partial p}(s_p^* - s_a^*) \cdot dp && (5.5) \\ &= \int_{T_n}^{T_i} (s_p^* - s_a^*) \cdot dT && (5.6) \end{aligned}$$

which is Equation (6.4.2) in Emanuel (1994).

Because the integration is now over temperature, and the difference between environment and parcel is taken at the same T, we can readily substitute saturated pseudo-enthalpy h^*

for the saturated pseudo-entropy s^* via:

$$\Delta h^* = T\Delta s* \tag{5.7}$$

Equation (5.6) then becomes:

$$CAPE = \int_{T_n}^{T_i} \frac{h_p^* - h_a^*}{T} dT \tag{5.8}$$

Here $h_p^* = h_s$ is conserved for an adiabatically rising parcel, while the moist static energy of the environment h_a^* is a weak function of T in individual atmospheric profiles, reaching a minimum in mid-troposphere that can be <15% below h_s.

Approximating the integral as a simple difference

All simplified frameworks for CAPE must replace the integral with some kind of simple difference. If the moist static energy difference between the parcel and the environment were independent of T, we could write

$$CAPE = (h_s - \overline{h^*}) \cdot ln\frac{T_i}{T_n} \tag{5.9}$$

This assumption is not perfect, and in practice, the true shape of atmospheric profiles necessitates adding an empirical coefficient to the relationship. Since an empirical coefficient is needed regardless, for convenience, we take the difference at the location of the minimum tropospheric MSE, typically around 650 mb.

$$CAPE \approx A \cdot (h_s - h_m^*) \tag{5.10}$$

which is the linear relationship used in this work; $(h_s - h_m^*)$ is the “ MSE surplus”. The coefficient A captures the shape of the profile, and, if virtual temperature corrections were indeed negligible, would be mathematically constrained to be between zero and $ln\frac{T_i}{T_n}$ [see also 2], at maximum ∼0.4 (for T_i = 300 K, T_n = 200 K). (In practice, compensating for the neglected virtual temperature corrections raises A slightly.) For the same temperature range, a larger A corresponds to a more uniform Δh profile between the lifting condensation level and the tropopause. For the dataset used in this work, the empirical slope A is 0.27.

Effect of assumptions in derivation

The derivation in E94 relies on two successive assumptions about the direct effect of water on the density of the environment through which a parcel rises. We show here that the effect of these assumptions is not prohibitive and is smaller than the effect of the core assumption in the alternative CAPE framework of Eq. (5) in [82, LC21]. The assumptions are:

- E94a: compute the virtual temperature effect for the environment assuming saturation
- E94b: neglect the virtual temperature effect for both parcel and environment
- LC21 Equation (5): assume all water vapor in the parcel condenses at the LCL

In reality, the mean environmental relative humidity in our high-CAPE midlatitudes summertime profiles is 0.44 (for all levels below 200 hPa). Both of the assumptions in E94 will therefore produce an underestimation of CAPE (the parcel is less buoyant than in reality), while that in LC21 will produce an overestimation (the parcel is more buoyant).

We illustrate the effects of these assumptions on an example atmospheric profile in Figure S11. The example profile is chosen to match the location and time of Figure 3 in [82]: Springfield, MO in early June. At 650 hPa, the true buoyancy $g\frac{\Delta T_v}{T_{ve}}$ is 0.143 m/s^2. The assumptions in E94 underestimate buoyancy by 14% and 22% (0.123 and 0.111 m/s^2), while that in LC21 overestimates it by a factor of 6 (0.845 m/s^2). The discrepancies are about half as large when averaged over the parcel's ascent but their relative sizes are unchanged: E94a,b cause underestimations of 6% and 7% and LC21 causes an overestimation of a factor of 3. If we use instead an average summertime profile over the Southeastern U.S., the bias produced by E94 remains below 13% while that in LC21 is a factor of eight. In both frameworks, the bias is largely accounted for by an empirical regression coefficient, but the more modest assumptions of E94 lead to a robust regression across climate states.

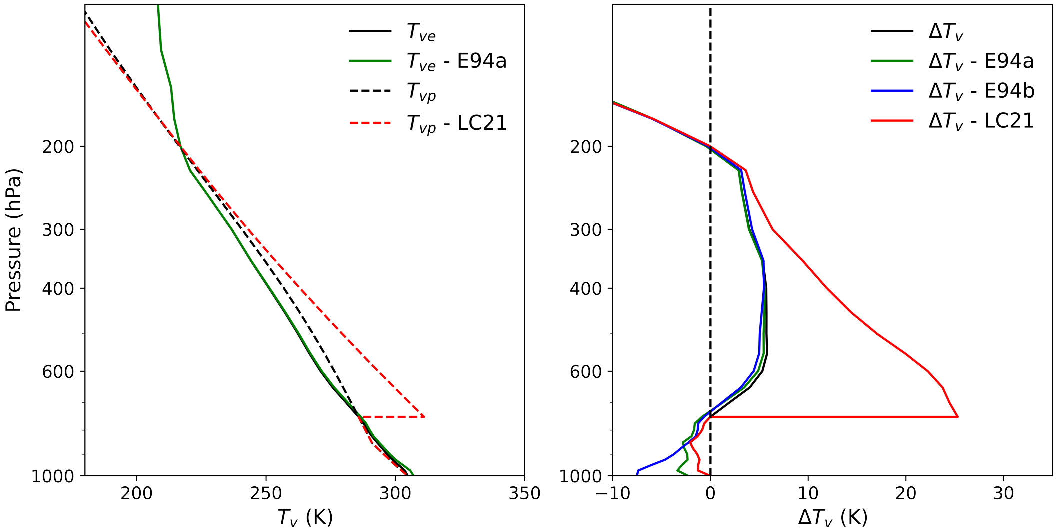

Figure S11: Illustration of the effect of assumptions in E94 and LC21, on (left) the virtual temperatures of parcel and environment and (right) the virtual temperature difference between parcel and environment. The example profile is that for a 1 × 1 deg grid in ERA5 including Springfield, MO, at 18 UTC on June 6th, 2005, chosen to approximately match the snapshot used in LC21 Figure 3 (0000 UTC June 07, 2011). E94a (green) raises the environmental T_v; E94b (blue) lowers T_v in both environment and parcel; and LC21 (red) raises T_v in the parcel by condensing all water at the LCL. The biases introduced by E94 are more modest than those in LC21, though all are ultimately accounted for by empirical regression coefficients.

S2.2 Model details

Table S4: List of CMIP6 Models included in the study. The outputs we use are 6-hourly model-level data for both *historical* and *ssp585* experiments. The data is available from ESGF.

Model	Variant label	Horizontal grid	Vertical levels
ACCESS-ESM1-5	r6i1p1f1	192 × 145	38
CanESM5	r1i1p2f1	128 × 64	49
CMCC-CM2-SR5	r1i1p1f1	288 × 192	30
CNRM-CM6-1	r1i1p1f2	256 × 128	91
CNRM-ESM2-1	r1i1p1f2	256 × 128	91
EC-Earth3	r1i1p1f1	512 × 256	91
GISS-E2-1-G	r1i1p1f2	144 × 90	40
MIROC-ES2L	r1i1p1f2	128 × 64	40
MPI-ESM1-2-LR	r1i1p1f1	192 × 96	47
MPI-ESM1-2-HR	r1i1p1f1	384 × 192	95
NorESM2-MM	r1i1p1f1	288 × 192	32

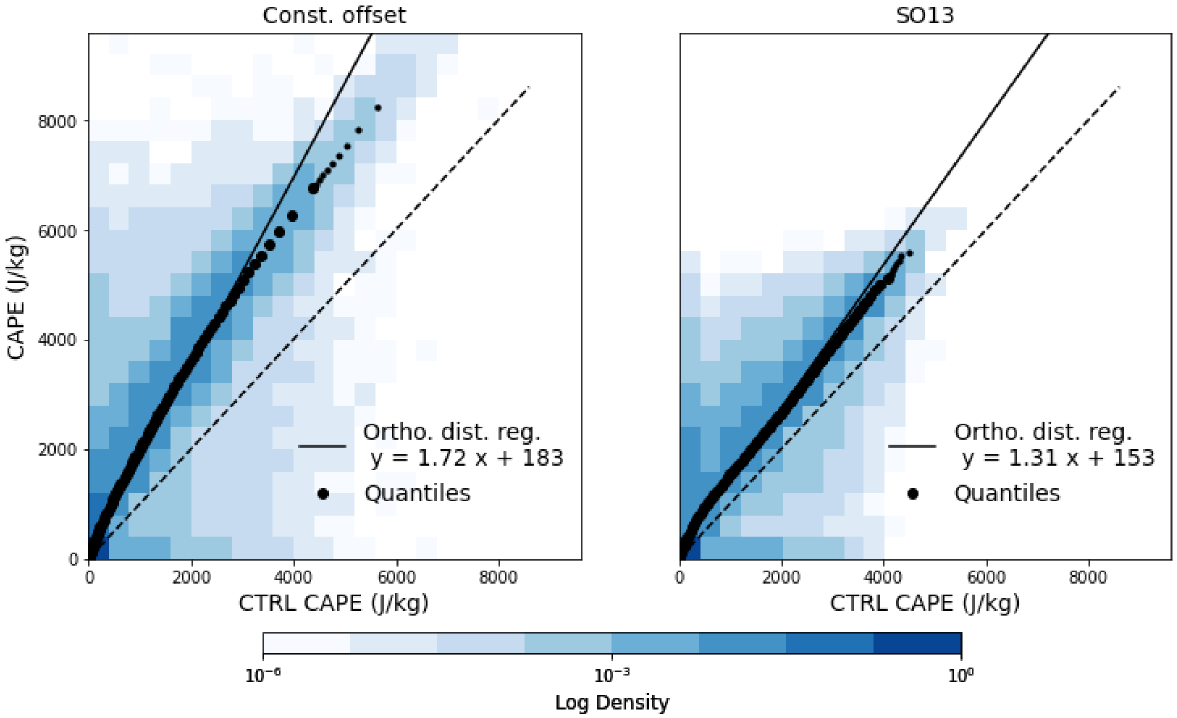

Figure S12: Changes between present and future CAPE, as in main text Figure 3.1 left panel, but for (left) *constant-offset* and (right) *SO13*, calculated as described in text. CAPE changes are too large in *constant offset* and too small in *SO13*: dividing by 4.65 K produces fractional changes of 12%/K and 6%/K, respectively, vs. the 8%/K derived from model output. For *constant offset* in particular, the quantiles fall below the orthogonal distance regression line above the 80th percentile. In both cases, however, the quantile regression matches the orthogonal distance regression reasonably well.

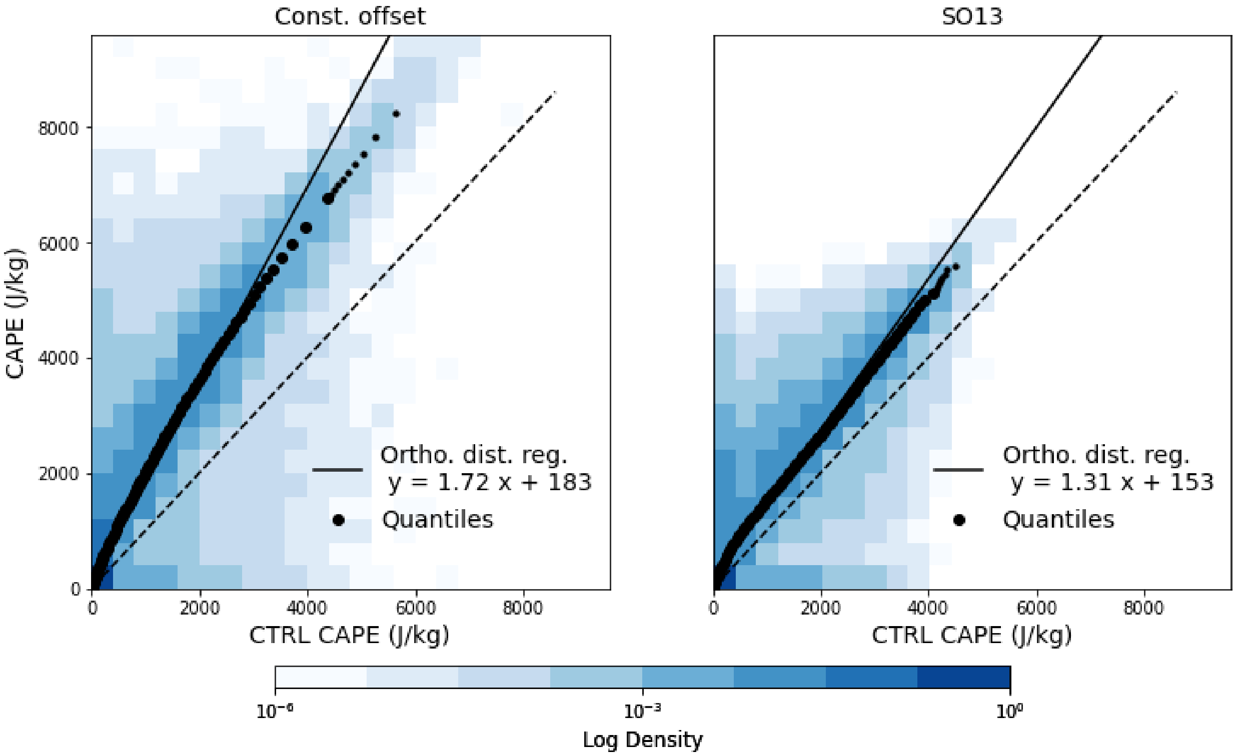

Figure S13: (Left and middle) Mean CAPE heatmap as in main text Figure 3.2 for present and future model output and (right) for the *constant offset* synthetic representation of future CAPE. Contours in black show CAPE of 2000 and 4000 J/kg in each panel, with CTRL contours repeated in gray in middle and right panels. Contours shift in PGW model output (center), meaning that warmer or wetter conditions are required to achieve the same CAPE. The *constant offset* synthetic (right), which involves changes in surface conditions alone and has no lapse rate adjustment, exhibits about half the shift of the PGW simulations.

Table S5: Evaluating synthetics: fitted slopes and intercepts of the future CAPE-MSE framework as in main text Figure 3.4, for actual PGW model output and for three synthetic datasets. *C–C* scaling produces too small a slope and *constant offset* too small an intercept. *Lapse rate adjustment* performs well at both.

	PGW	C–C	Constant offset	Lapse rate adj.
slope	0.239	0.271	0.240	0.236
x-intercept (kJ/kg)	346.2	350.4	343.8	345.8

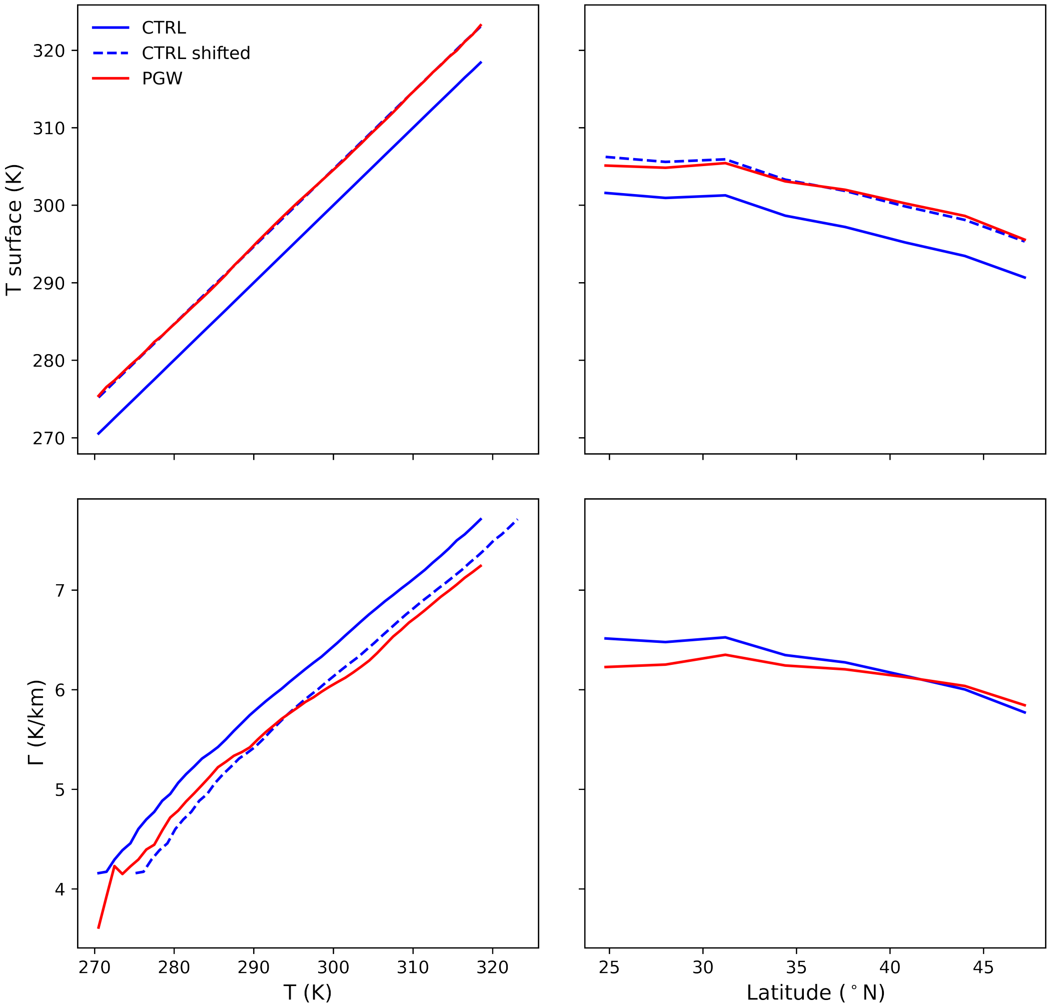

Figure S14: Changes in (top) temperature and (bottom) lapse rate as a function of (left) temperature bins and (right) latitudinal bins. The blue dashed lines are synthetics applying a same 4.65 K offset to the CTRL climate (blue solid lines). The damping of CAPE under a warmer climate can be explained by a more stable lapse rate associated with a given set of surface conditions.

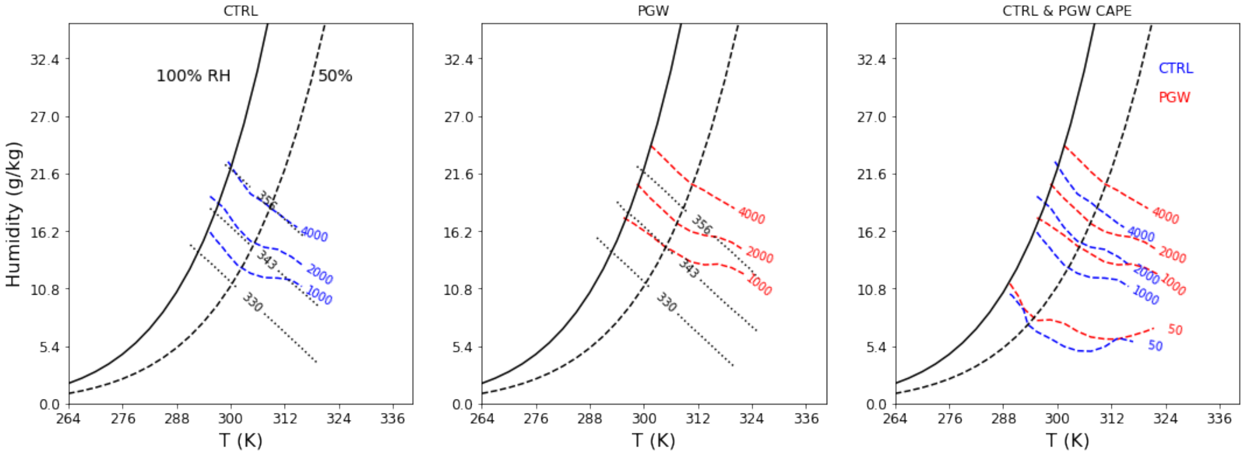

Figure S15: Contours of CAPE and surface moist static energy (MSE) in model output for simulations in present (CTRL, left) and future (PGW, middle) conditions. CAPE contours follow those of moist static energy in the convection-promoting regime (CAPE > 1000 J/kg, RH > 40%), The relationship differs between CTRL and PGW (right). Contours here are cut off at RH=100%, as in main text Figure 3.2. CAPE contours aligns with those of surface MSE in conditions with high CAPE, suggesting a strong dependence between CAPE and surface MSE. Future changes in CAPE can be translated to a change in mapping between surface MSE and CAPE.

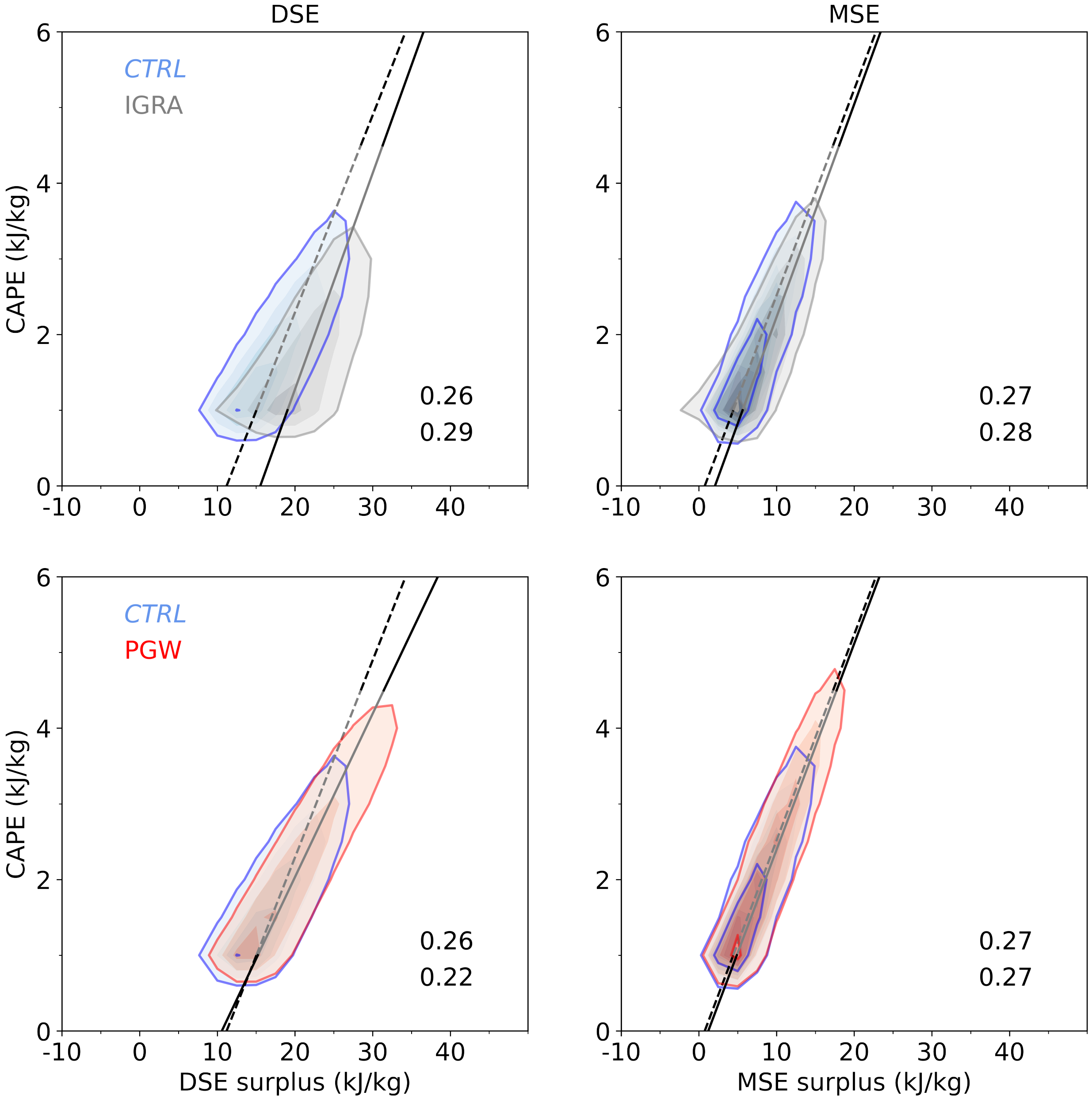

Figure S16: Comparison of the CAPE relationship with dry static energy (DSE) surplus (left) and with MSE surplus (right). DSE surplus is defined as the difference between surface MSE and mean mid-tropospheric DSE (virtual-temperature weighted free troposphere DSE). The top rows show model output (WRF) versus observations (IGRA) under CTRL climate, and the bottom rows show CTRL versus PGW in model output. Color shading increments are 1.5% for all panels, and the text shows the slopes for CTRL and IGRA/PGW. Conclusions are 1) the WRF simulation realistically reproduces the observed joint distribution of CAPE and MSE surplus and 2) a linear expression with MSE surplus outperforms that with DSE surplus both in residual variance and in robustness.

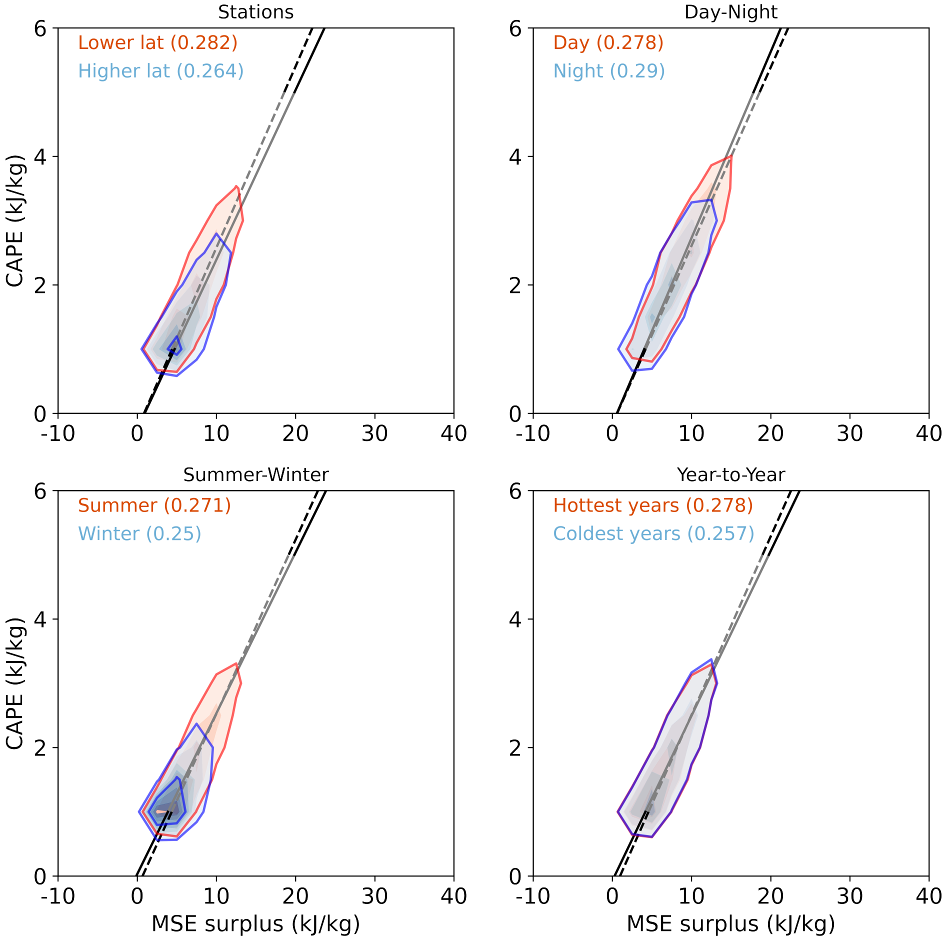

Figure S17: Tests of the robustness of the CAPE-MSE surplus relationship with different subsets of the data. *Top left*: stations lower and higher latitude than 35N. *Top right*: daytime versus nighttime (using only stations below 30N, to avoid biasing the sampling). *Bottom left*: summertime (MJJA) versus wintertime (NDJF) (all other panels use summertime data only; note that the month of February 2005 in the PGW run is removed due to missing surface 2D fields). *Bottom right*: the hottest 3 years (2001, 2006, 2012) versus the coldest 3 years (2004, 2008, 2009). Figure uses only CAPE $\geq$ 1000 J/kg, and all panels besides lower L. use summertime data only. MSE surplus is derived using the minimum saturation MSE in each individual profile. Each color shading is a 1.5% increment in density, and the orthogonal regression is fit using binned median values. The CAPE-MSE surplus framework (including its intercepts and slopes) is highly consistent across all cases tested.

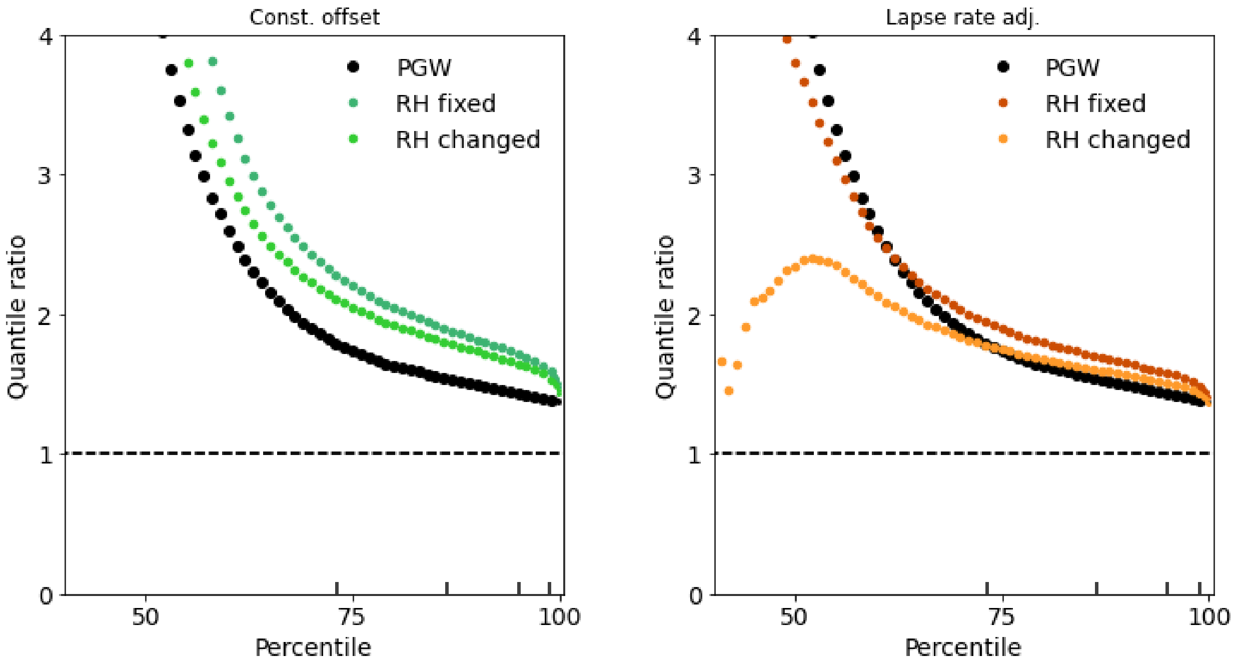

Figure S18: Quantile ratio plots of future–present CAPE in model output (PGW vs. CTRL) and in synthetic future distributions, showing also the effect of RH changes. For each synthetic we show one version with constant surface RH and one with a uniform ∼1% reduction, which lowers future CAPE changes by about 6% in both cases. (Left) *Constant offset.* Mean fractional changes are 1.92 with fixed RH and 1.81 with the reduction. Values are derived from the average quantile ratios for ≥ 73rd percentile. (Right) *Lapse rate adjustment.* Mean fractional changes are 1.71 and 1.61.

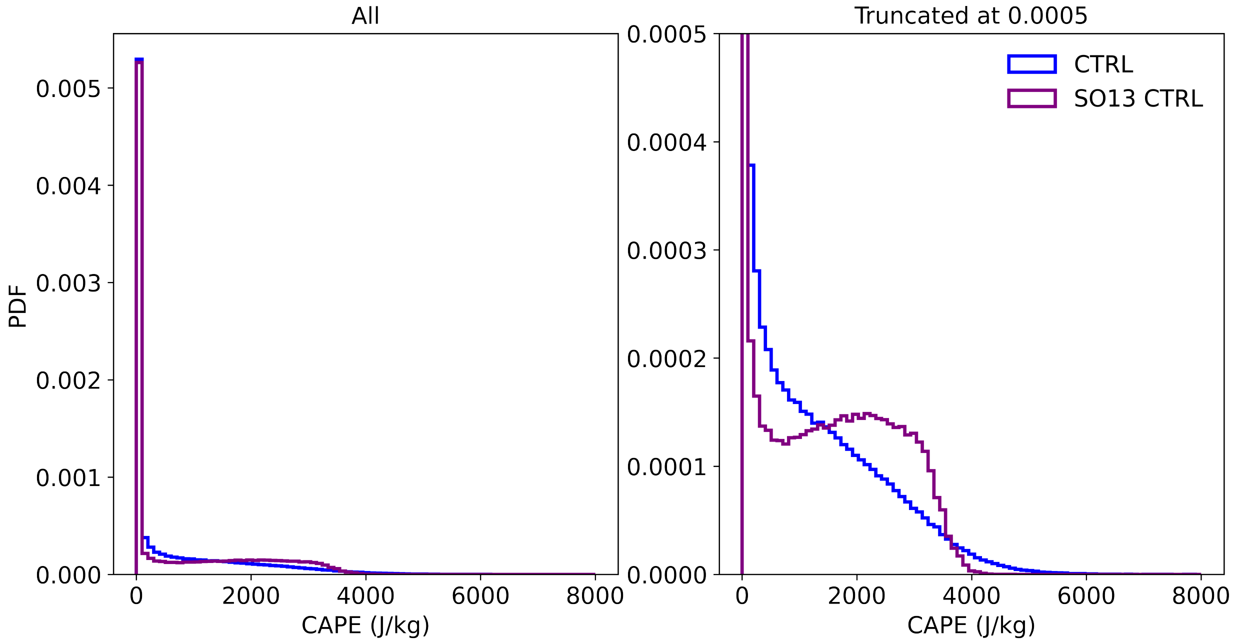

Figure S19: Comparison of probability distributions of midlatitudes summertime CAPE under current CTRL climate in our WRF model output and in the SO13 zero-buoyancy model driven by CTRL surface temperatures. (Left) full CAPE distribution and (right) with y-axis truncated at 0.0005 to show detail. The zero-buoyancy model cannot reproduce a realistic distribution, overestimating the occurrence of moderately high CAPE (between 1500 and 3500 J/kg). This peak occurs even though we modify the SO13 procedure to force the integration to stop at the actual LNB of each profile. Without that modification, SO13 cannot produce zero-CAPE values and the distribution is even more unimodal. SO13 is designed to reproduce climatological means in strongly convecting regions and is not appropriate for midlatitudes land.

S3 Supplemental Material for Chapter 4

S3.1 Derivation of full CAPE scaling

Here we show the full derivation of the CAPE scaling framework – the goal is to develop a simple scaling for CAPE that works across space and different time scales without the need to carry out integration.

Derivation based on entropy form

We start from the expression in Emanuel (1994) Eq. (6.4.2):

$$CAPE = \int (s^*_{pp} - s^*_{pa}) dT \tag{5.11}$$

here we assume that the "effect of water content on density can be neglected". We will show the reasoning for the validity of such an assumption in the next section ("Derivation based on buoyancy form"). We will denote the environment with subscript e instead a for consistency.

Since we are integrating at the same temperature, we could convert the entropy to Moist Static Energy (or MSE) with $\Delta s \approx \frac{\Delta h}{T}$ (assuming the temperature difference between the parcel and the environment is negligible). Then we have:

$$CAPE = \int \frac{h^*_p - h^*_e}{T} dT \tag{5.12}$$

where $h_p = h^*_p = h_s$ is the surface MSE, conserved for parcel rising adiabatically. Approximating both lapse rate and environmental temperature by their mean values, we have:

$$CAPE = \int \frac{h_p - h^*_e}{T} \Gamma \, dz \tag{5.13}$$

Since the saturation MSE difference is a function of z (or T), and it is always below the

MSE surplus (Δh, defined as the maximum difference between surface MSE and saturated environmental MSE), we could replace the integral with a "shape of profile" parameter ($\alpha(z) = \frac{h_p - h_e^*}{\Delta h}$), which is always constrained between 0 and 1; we further define β as the factor $\frac{\Gamma}{T}$.

Further simplifying β and α to a single scaling factor $\hat{k} = \frac{\int \beta\alpha dz}{\Delta z}$, we have the final form of CAPE scaling:

$$CAPE = \hat{k}\Delta h\Delta z \approx k\frac{\overline{\Gamma}}{\overline{T}}\Delta h\Delta z \tag{5.14}$$

where Δz ($=z_{LNB} - z_{LFC}$) is the convective layer depth. Throughout this work, we assume that both lapse rate and column mean temperature could be approximated with their mean value on the spatial and temporal scale evaluated.

Note that the "shape of profile" parameter α is mathematically constrained to a number close to 0.5 for most models, moderate to high CAPE profiles. If the whole profile (weighted by depth for each layer) is rectangular, α would be twice as large as a triangular-shaped profile. This number is empirically fitted for every validation we carry out in the manuscript and is found to be largely constant across different sub-regions, between climate states, and between summer and winter. The assumption of a constant k starts to break down if the dryness varies substantially across space and time scale of interest, and the buoyancy factor contributes to a non-negligible deviation from a fixed constant. Nevertheless, with a mathematically constrained free parameter, our scaling framework shows robustness across almost all scales and all models evaluated.

Derivation based on buoyancy form

The scaling could also be derived from the buoyancy form, a more widely known CAPE definition:

$$CAPE = \int g\frac{T_{vp} - T_{ve}}{T_{ve}}dz \tag{5.15}$$

To derive a full scaling, a strong assumption that the water vapor effect on density is negligible has to be made. This is consistent with the assumption made by Emanuel (1994) in the derivation of the entropy form of CAPE. With this assumption, we replace the virtual temperature with the raw temperature in Eq. 5.15. It is noted in Emanuel (1994) that this error could be as large as 100% for certain profiles in the deep tropics (dry environment but very moist surface parcel). With this assumption, we approximate CAPE as:

$$CAPE = \int g \frac{T_p - T_e}{T_e} dz \tag{5.16}$$

Given that we are evaluating buoyancy at the same height above the level of free convection, the difference in the geopotential term disappears. Since the buoyancy is mainly a function of temperature, the difference in T can be linearly mapped to the difference in saturation MSE between the parcel and environment using the Clausius-Clapeyron relationship:

$$h_p - h_e^* = h_p^* - h_e^* = c_p(T_p - T_e) + \frac{L_v^2}{R_v T_e^2} q^*(T_p - T_e) \tag{5.17}$$

Substituting into the CAPE expression, we have:

$$CAPE = \int \frac{g}{c_p(1 + \frac{L_v^2}{R_v T_e^2 c_p} q_e^*)} \frac{h_p - h_e^*}{T_e} dz \tag{5.18}$$

Given that moist adiabatic lapse rate taking the form of $\Gamma_m = \frac{g}{c_p} \frac{1+\frac{Lvq_e^*}{R_d T_e}}{1+\frac{L_v^2 q_e^*}{R_v T^2 c_p}}$, we have:

$$CAPE = \int \frac{\Gamma_m}{1 + \frac{L_v q_e^*}{R_d T_e}} \frac{h_p - h_e^*}{T_e} dz \tag{5.19}$$

The second terms on the numerator and denominator are $O(10^{-1})$ of the first term T_e. If we were to neglect variations in the scaling factor, we have a simplified buoyancy determined by T alone:

$$b_T = \Gamma_m \frac{h_p - h_e^*}{T_e} \tag{5.20}$$

The remaining factor $\frac{1}{1+\frac{L_v q_e^*}{R_d T_e}}$ directly modifies the buoyancy at each level:

$$b = \frac{1}{1 + \frac{L_v q_e^*}{R_d T_e}} b_T \tag{5.21}$$

The remaining factor reduces the buoyancy, and its effect is larger at hotter temperatures. (This effect never exceeds 50%, even at a hot surface temperature of around 305 K.) But it is a pure dependence on T, so for a given temperature, the factor remains largely unchanged. This suggests that its integrated value will also remain relatively constant with a warming of a few degrees (see Figure S30 for details).

To reduce the integration to a simplified scaling, we replace $h_p - h_e^*$ with $(h_p - h_e^*)_{max} \equiv \Delta h$.

$$CAPE = \int \beta (h_p - h_e^*) dz = \Delta h \int \beta \alpha dz \tag{5.22}$$

where β is the buoyancy factor multiplied by $\frac{\Gamma_m}{T_e}$, and the "shape of profile" parameter $\alpha(z) = \frac{h_p - h_e^*}{\Delta h}$.

Further reducing β to $\hat{k} = \frac{\int \beta \alpha dz}{\Delta z}$, we have the final CAPE scaling:

$$CAPE = \hat{k} \Delta h \Delta z \approx k \frac{\overline{\Gamma}}{\overline{T}} \Delta h \Delta z \tag{5.23}$$

Given that we only need a scalar for Γ for the analysis shown in the main text, here we approximate Γ_m with column mean environmental lapse rate.

S3.2 Regression methodology

In this work, two types of regressions are carried out:

Ordinary Least Square (OLS) is used when fitting the “shape of profile” parameter k given all other terms in main text Eq. (2). We use the simplest model without intercept because CAPE should develop when the MSE surplus exceeds zero. Only cases with CAPE >100 J/kg are considered in these regressions to emphasize the moderate to high CAPE conditions.

Orthogonal Distance Regression (ODR) is used when fitting the fractional changes between the same variable across two states (current and future, night and day, or winter and summer). The regression is selected to minimize errors in both dependent and independent variables. The slope is the ratio between the two states, i.e., a mean change that best describes the shift in joint distribution space.

S3.3 2-sample K–S test

We use a two-sample Kolmogorov–Smirnov (KS) test to determine whether the distributions of modeled CAPE and our scaling can be obtained by drawing from the same distribution. The null hypothesis is that the distributions of the two samples are the same. We reject the null hypothesis at confidence level $(1-\alpha)\times 100\%$ when the test statistic $D_{n,m}$ satisfies

$$D_{n,m} > c(\alpha)\sqrt{\frac{n+m}{n\cdot m}}, \quad c(\alpha) = \sqrt{-ln(\frac{\alpha}{2})\cdot\frac{1}{2}}$$

where $D_{n,m}$ is the maximum difference between the two empirical cumulative distribution functions and n and m are the number of sample points in the two samples being compared (e.g. n for ERA5 is $\sim$ 8 million). In this analysis, we have chosen $\alpha = 0.05$ to test at the 95% confidence level.

S3.4 Regional “shape of profile” parameter derivation

The “shape of profile” parameter k can be assumed as a constant on certain regional spatial scales, but its variation can be significant for locations with substantially different humidity (e.g., land-sea contrast). Therefore, we must derive regional k to construct our scaling across space properly.

To construct the scaling, we use convoluted boxes of 21 x 21 degree sizes (see Figure S24 for performance), and calculate averages of all relevant physical parameters in these boxes. A regional k could be obtained from main text Eq. (2) based on convoluted mean parameters, and we reconstruct regional scaling by multiplying regional k to physical parameters at native resolution.

S3.5 Supporting Figures

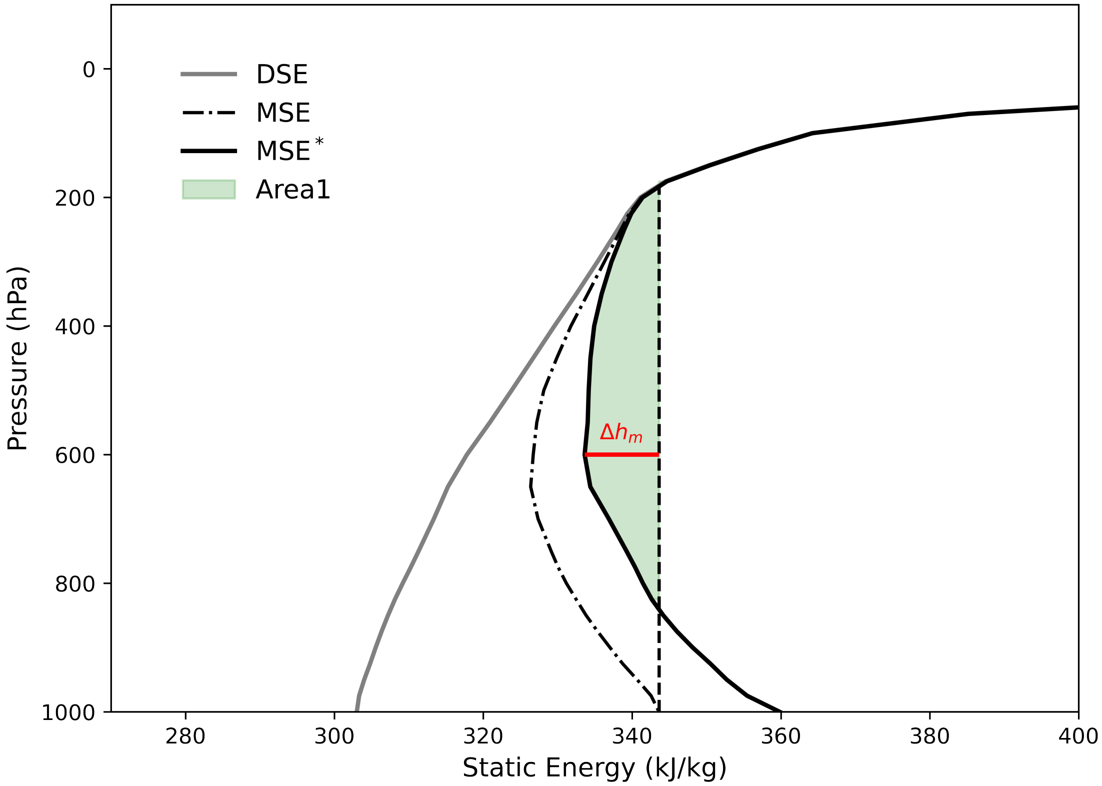

Figure S20: Vertical profiles of continental U.S. for Dry Static Energy (DSE, gray solid), Moist Static Energy (MSE, black dot-dashed), and saturation MSE (black solid), during a high CAPE episode from June 6th–8th, 2005. The rising parcel conserves MSE (black dashed) as it ascends adiabatically. The red line corresponds to MSE surplus, defined as the maximum difference between saturation MSE (env.) and surface MSE (parcel), and Area 1 (green shading) corresponds to the integrated difference between them. CAPE is proportional to Area 1 under the assumptions stated in the derivation above.

Table S6: Percentage of total variance explained by each term and some combinations in the full scaling in ERA5 in current climate, Eastern U.S.

Varying term	Δh	Δz	$\overline{T}$	$\overline{\Gamma}$	$\Delta h + \Delta z$	$\overline{T} + \overline{\Gamma}$
Interannual	25%	16%	0.13%	1%	76%	2%
Seasonal	19%	20%	0.03%	0.2%	90%	0.3%
Diurnal	21%	25%	0.02%	0.3%	88%	0.4%
Spatial	57%	7%	0.03%	0.8%	106%	1%

Table S7: Percentage of variance explained by each term and some combinations in the full scaling in ERA5 in current climate across the globe.

Varying term	Δh	Δz	$\overline{T}$	$\overline{\Gamma}$	$\Delta h + \Delta z$	$\overline{T} + \overline{\Gamma}$
Interannual	15%	29%	0.4%	5.4%	69%	8%
Seasonal	32%	23%	0.04%	0.4%	99.6%	0.7%
Diurnal	52%	17%	0%	0.09%	103%	0.1%
Spatial	15%	16%	0.03%	0.2%	83%	0.4%

Table S8: Summary of mean CAPE(J/kg), mean bias (J/kg) relative to ERA5, and "shape of profile" parameter fitted for each model. The full dataset (6-hourly data) is used without aggregation.

Model/Reanalysis	Mean CAPE	Mean bias	Fitted k
ERA5	389	/	0.516
ACCESS-ESM1-5	406	17	0.524
CanESM5	363	-26	0.478
CMCC-CM2-SR5	519	130	0.508
CNRM-CM6-1	759	370	0.599
CNRM-ESM2-1	782	393	0.599
EC-Earth3	380	-9	0.501
GISS-E2-1-G	873	484	0.592
MIROC-ES2L	567	178	0.504
MPI-ESM1-2-LR	239	-150	0.466
MPI-ESM1-2-HR	235	-154	0.480
NorESM2-MM	429	40	0.497

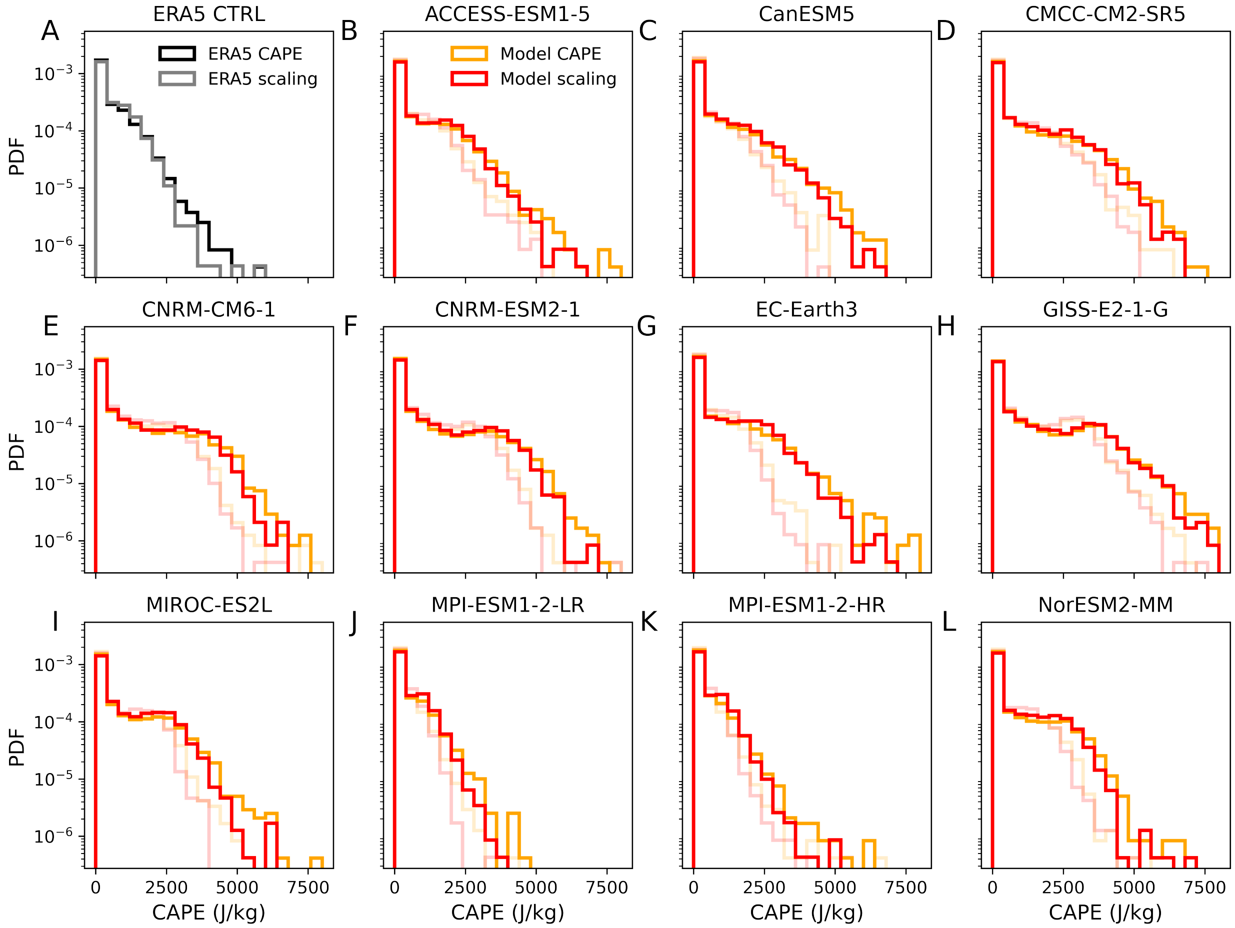

Figure S21: Probability density functions of CAPE (black/red) and scaling (gray/orange) in ssp585 runs (2091–2100) against historical runs (2005–2014), for (top left) ERA5 historical period and (other panels) 11 CMIP6 models. Color coding are the same main text Figure 4.1; ERA5 historical period are shown with thick lines, and CMIP6 models with fainted lines. The future CAPE distribution are well captured by the scaling across models.

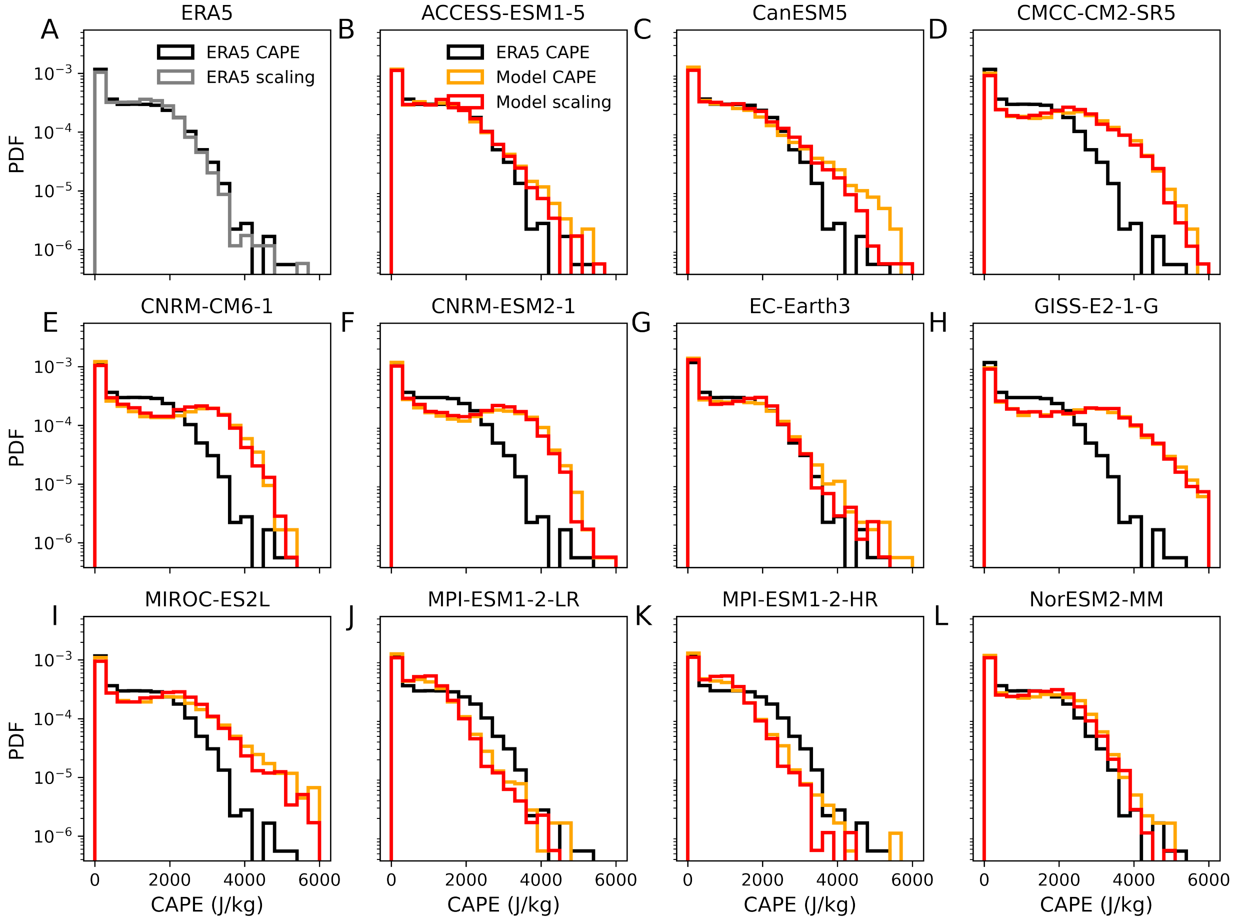

Figure S22: Probability density functions for reanalyses and all models as in main text Figure 4.1, but for Contiguous U.S. The performance is even better for a mid-latitude land sub-region in terms of fractional bias.

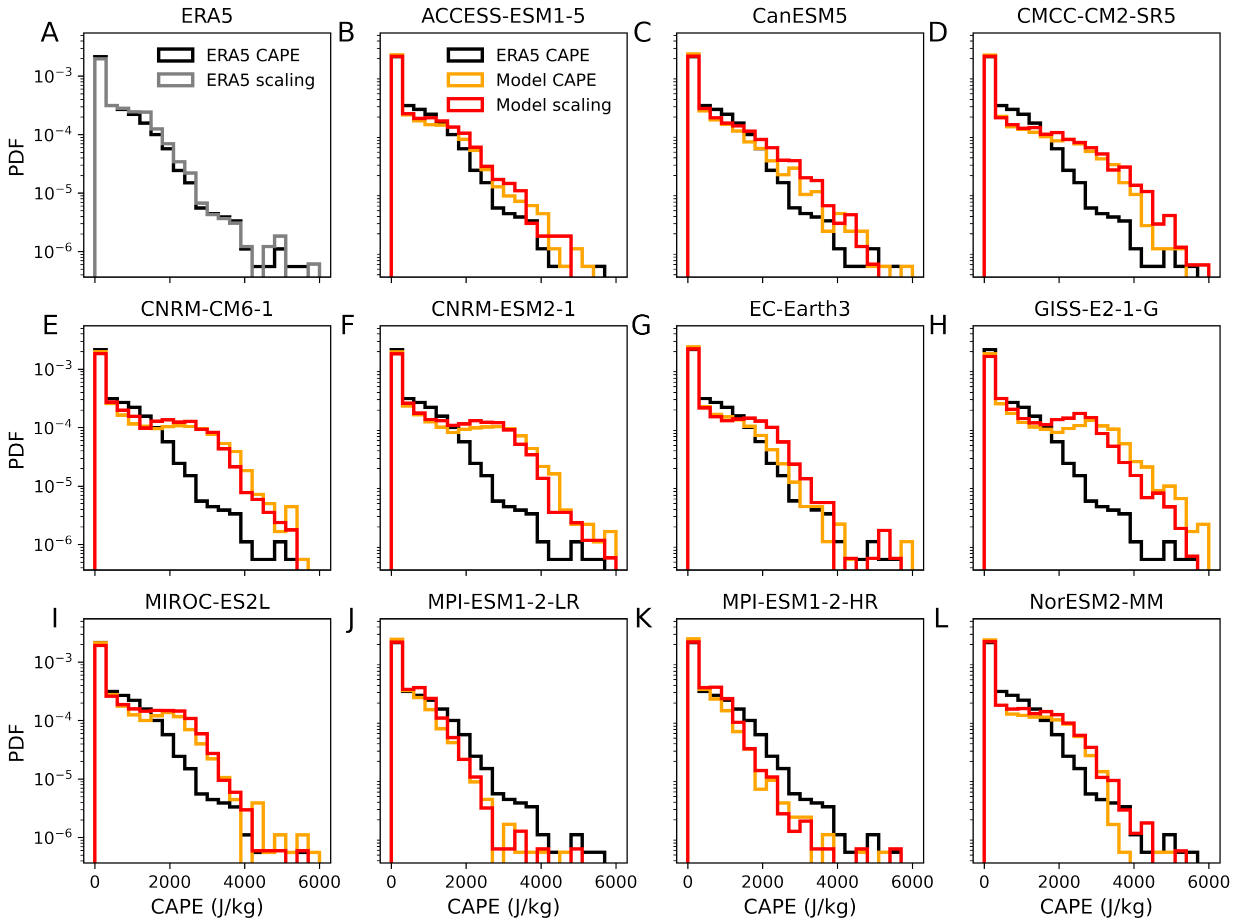

Figure S23: Probability density functions for reanalyses and all models as in main text Figure 4.1, but for scaling based on a global constant k. The global scaling performs extremely well in reproducing the modeled CAPE.

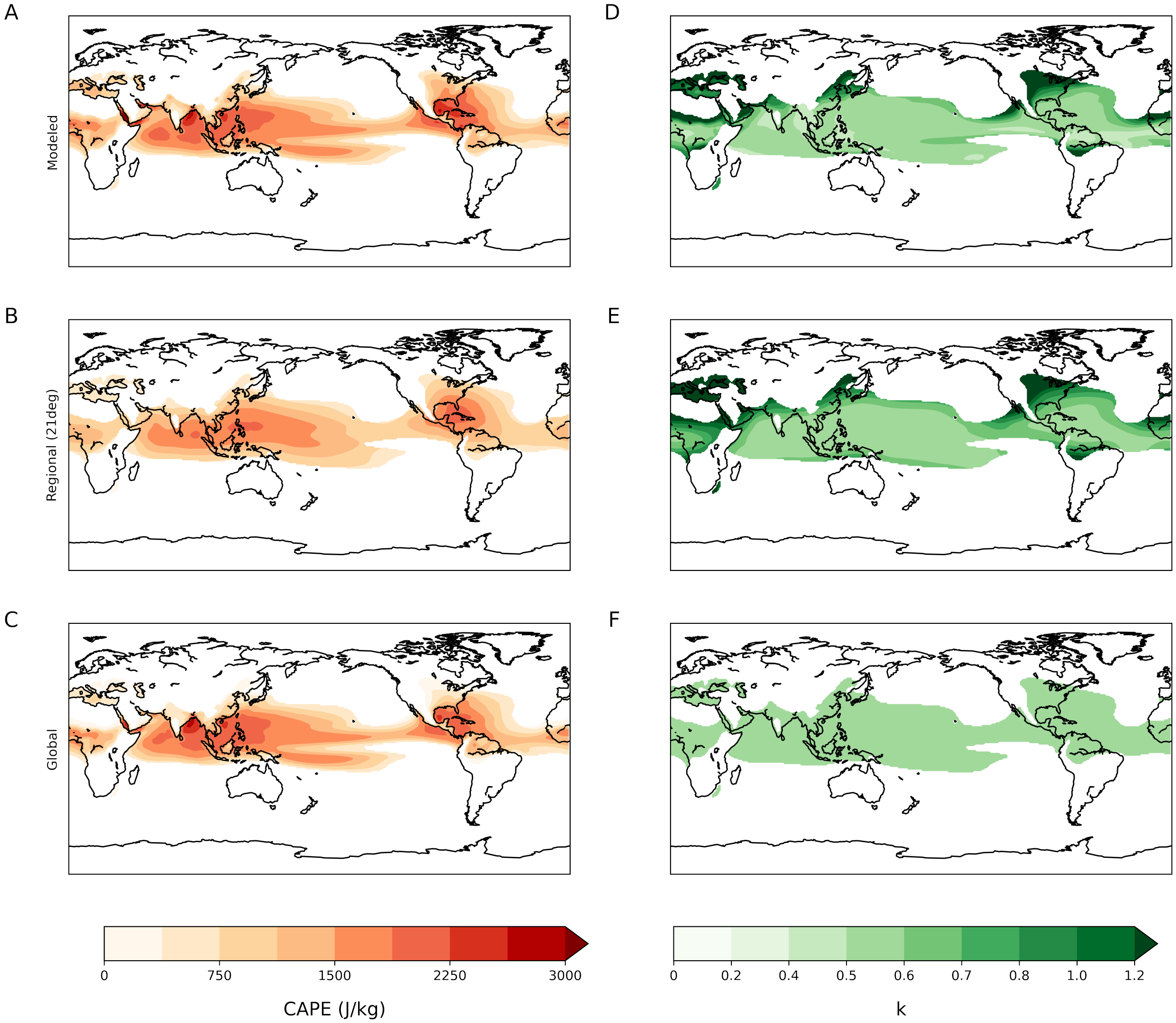

Figure S24: (Left) Mean CAPE scaling and (right) “shape of profile” k derived at different spatial grid resolutions. (Top) modeled CAPE and true k at native model resolution, (middle) regional scaling and k derived from 21 x 21 deg convolution means, (bottom) scaling and k assuming a global constant k applies. The scaling is able to capture CAPE variations over the tropical ocean with a global constant k, but performance over midlatitudes land has substantially improved.

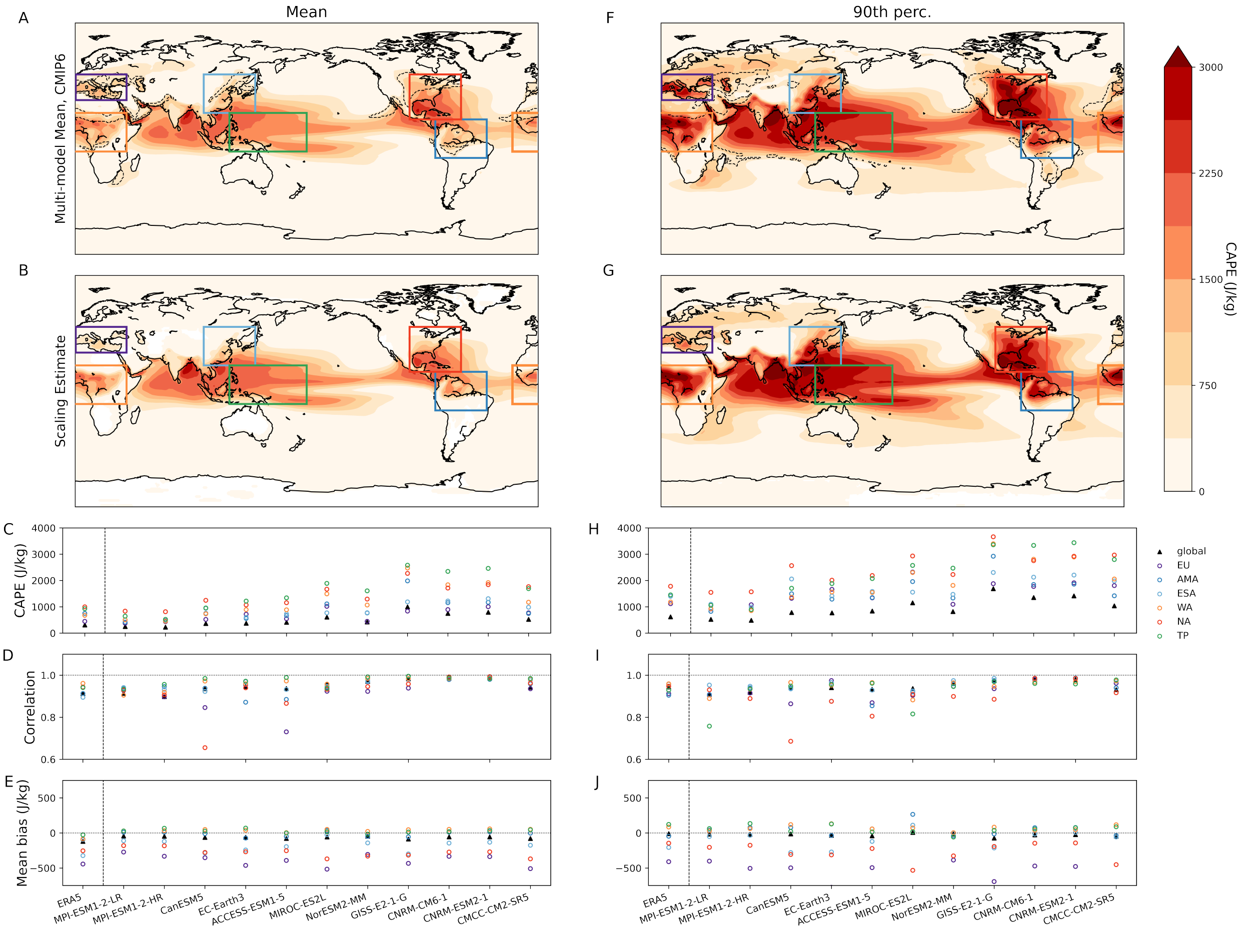

Figure S25: Spatial evaluation as in Figure 4.2, but for scaling based on a global constant k. Note the mean CAPE differs from Figure 4.2 because we are using raw CAPE from each model and not implementing any convolutional mean or interpolation. The scaling with a global constant k sufficiently captures spatial variations of CAPE over the tropical ocean, but severely underestimate mean CAPE over midlatitudes land surface.

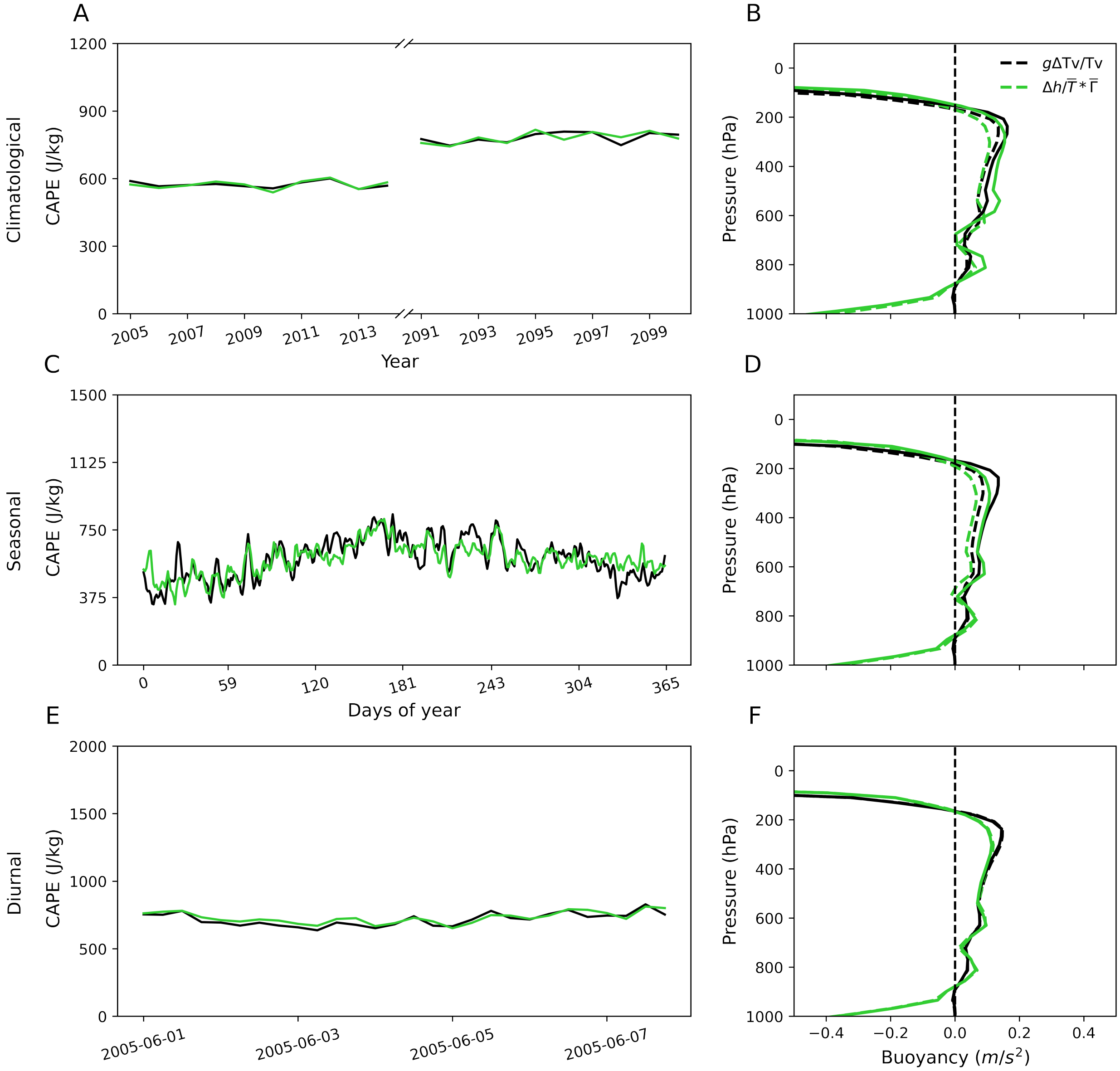

Figure S26: Time series and profiles as in main text Figure 4.3, but for Tropical Pacific. The diurnal variation magnitude is substantially smaller over tropical oceans, but the scaling still captures most of the temporal variations and mean profiles.

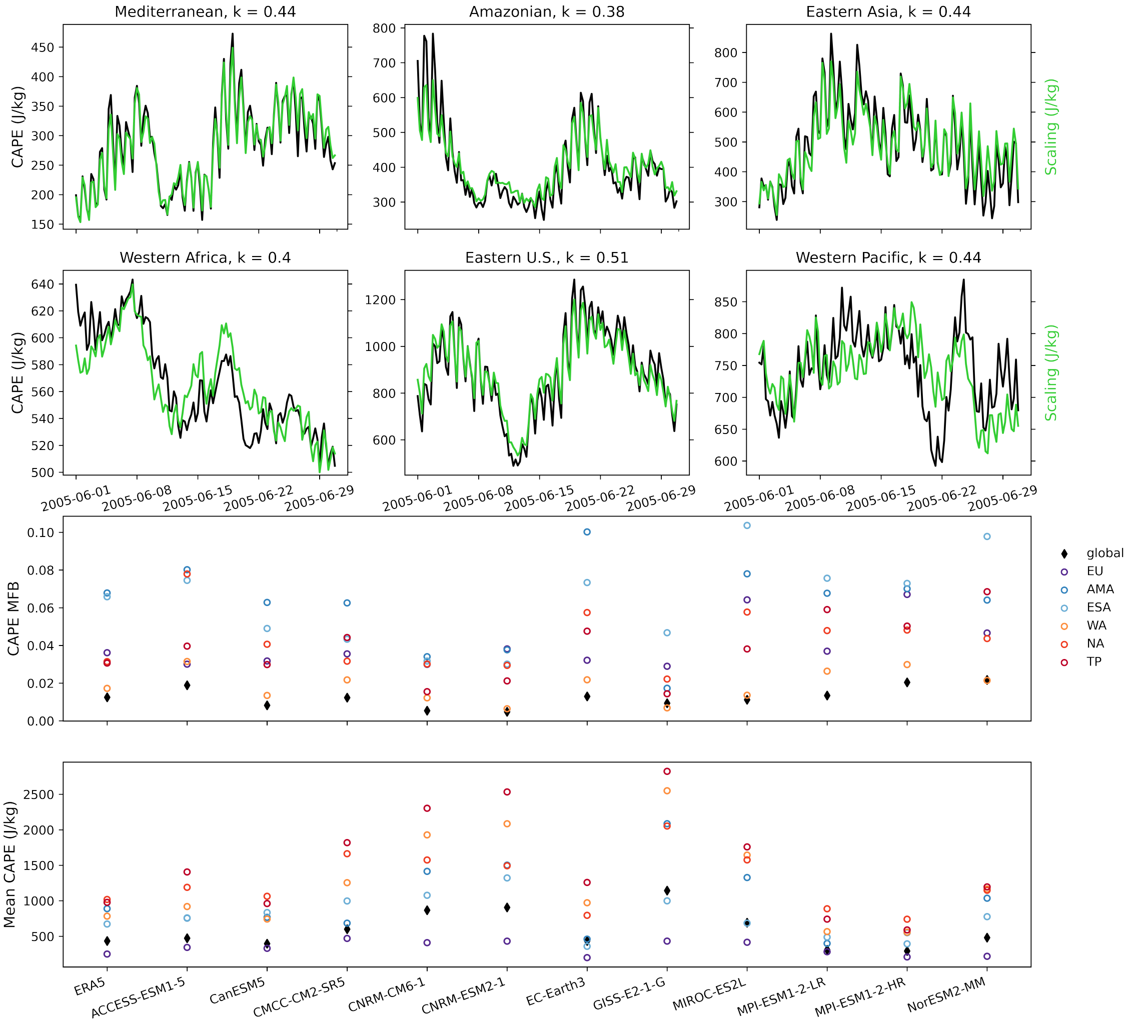

Figure S27: (A-F) Time series of CAPE and CAPE scaling across different sub-regions in MPI-ESM1-2-LR in June 2005, and evaluation of (G) mean fractional bias and (H) mean CAPE in the period shown above, across models (x-axis). Both the diurnal variation and episodes of highs and lows in CAPE are well-captured by our scaling.

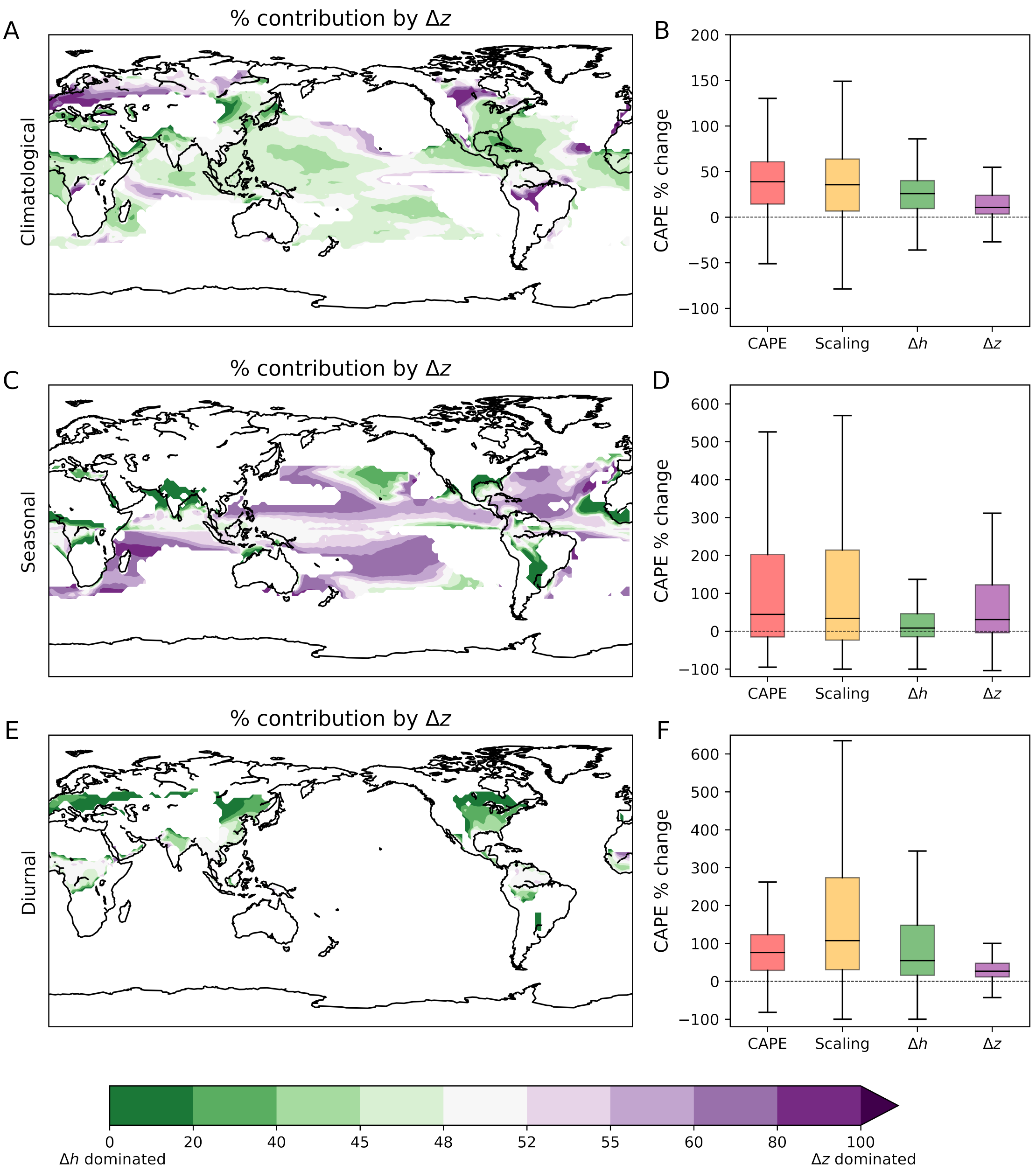

Figure S28: CAPE fractional changes on (A, B) climatological (future/control, JJA), (C, D) seasonal (summer/winter), and (E, F) diurnal (day/night, JJA) time scale, for MPI-ESM1-2-LR model. (Left) Spatial pattern of decomposed contribution to CAPE fractional changes from two parameters: A green-ish color means the majority of changes are attributed to changes in MSE surplus (Δh), while a blue-ish color means changes are dominated by convective layer depth (Δz). Only locations with CAPE $>$ 100 J/kg in the reference state are shown. (Right) Boxplots showing the spread of fractional change across space for modeled CAPE, scaling, and contribution from two decomposed parameters. The decomposition into MSE surplus and convective layer depth shows different parameters dominate over different regions on different time scales; the scaling captures the spread of the modeled CAPE changes pretty well.

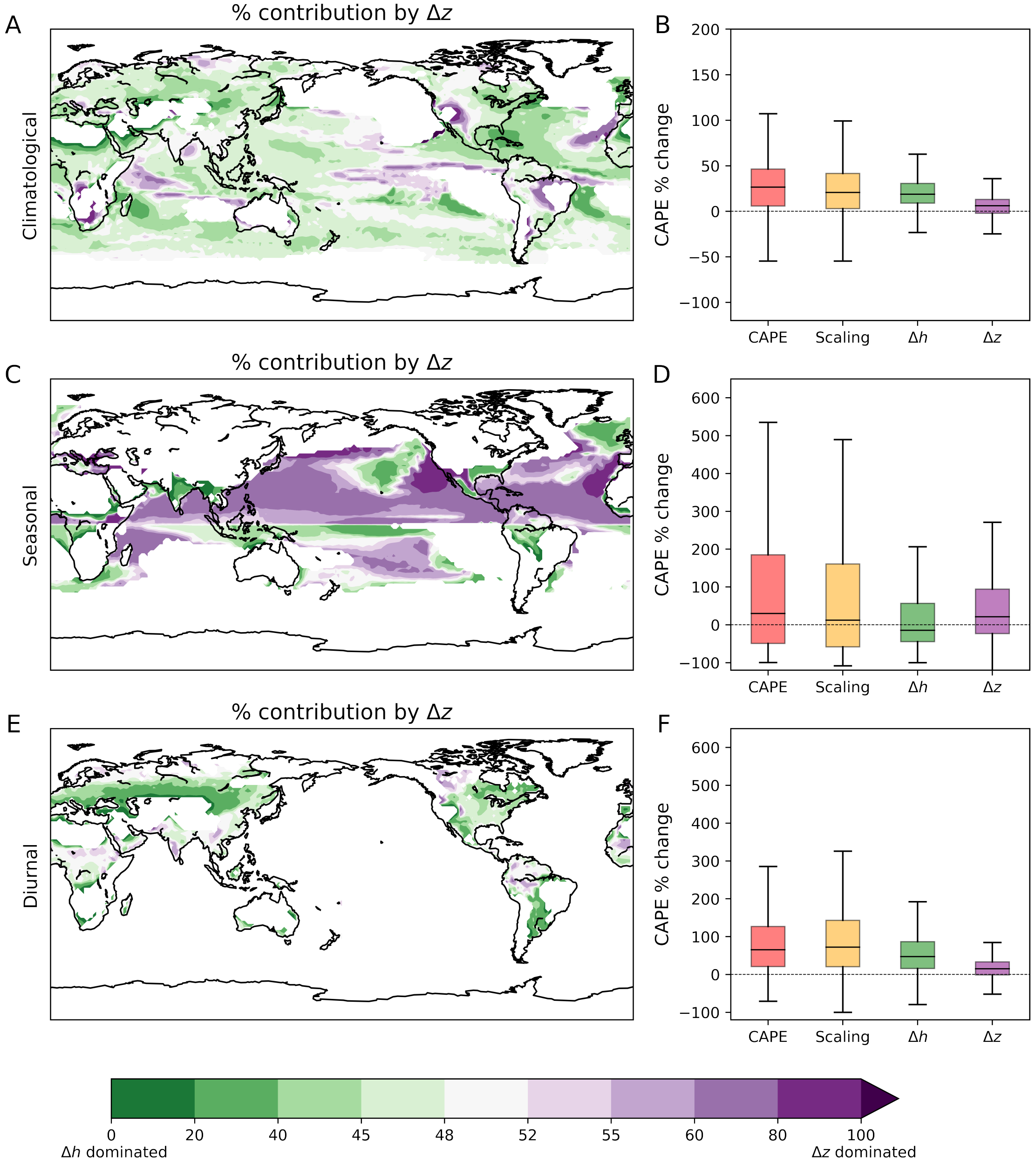

Figure S29: Decomposition of CAPE variations and dominant factors for 90th percentile, instead of mean as in Figure S28. The rows from the top are climatological, seasonal, and diurnal time scales, respectively. The climatological and diurnal CAPE variations are driven by MSE surplus, while the seasonal variations are driven by convective layer depth. These findings hold for both mean and extreme levels of CAPE.

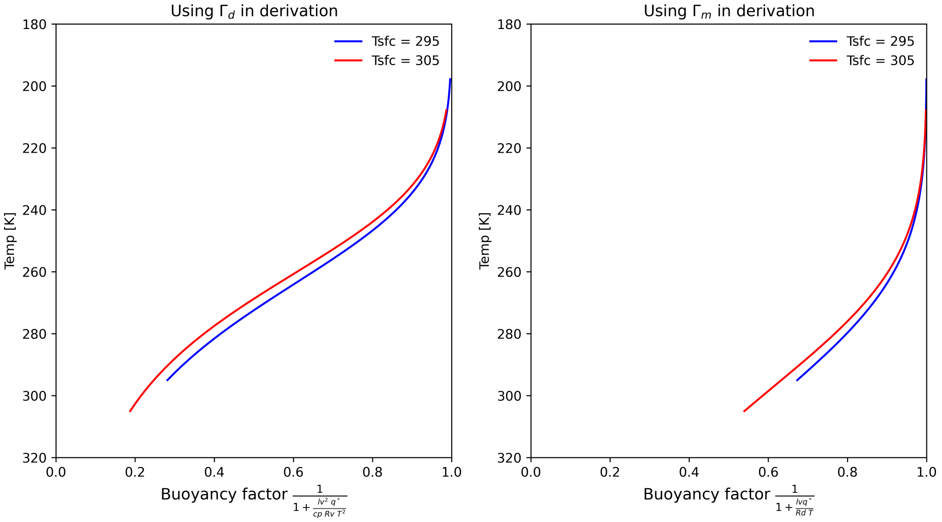

Figure S30: Buoyancy factors in CAPE scaling using (left) dry adiabatic lapse rate or (right) moist adiabatic lapse rate in derivation. We assume the vertical stratification follows a constant lapse rate at 6.5 K/km, and the pressure follows an exponential falloff determined by scale height. The blue and red lines correspond to surface temperatures of 295 and 305 K, respectively. Using a moist adiabatic lapse rate absorbs part of the T dependence and results in a more constant buoyancy factor across different vertical levels. The integral of the buoyancy factor remains largely unchanged with a few degrees of warming.